Credit Risk in the Context of European Integration: Assessing the Possibility of Pan-European Scoring

Galina Andreeva

Submitted for the degree of Doctor of Philosophy

The University of Edinburgh

2004

Statement of original authorship

Hereby I declare that this thesis has been composed by myself, the work is my own and it has not been submitted for any other degree or professional qualification, except as specified on the title page.

Galina Andreeva

February, 2004

Abstract

Credit scoring is the collection of techniques used for risk assessment in consumer credit. Traditionally a credit scoring model is constructed to fit a specific credit portfolio which normally consists of residents of one country (customised models). But the political desire for further integration of the European Union into the single internal market opens the possibility for the lenders to compete across national borders. Therefore the necessity arises to assess the risk of a mixed heterogeneous population consisting of residents of several European countries.

This thesis shows how a single generic model can be used to credit score the applicants for a revolving store card from three different European countries. First, the EU harmonisation process is reviewed with the aim to establish its likely impact on credit scoring practice. In particular, the legal restrictions on the information used in credit scoring models are examined and the effect of such restrictions for both lenders and borrowers is investigated. A comparison of credit regulations is provided between the USA and the EU, and for the latter the differences in the national legislation of the EU member states are presented.

Second, several generic models are developed using logistic regression and survival analysis, and their predictive accuracy is benchmarked against the performance of equivalent national (customised) models. Whilst logistic regression is the most established approach in the credit industry, survival analysis is a relatively new application that offers an advantage of predicting time to the event of interest and therefore, lays the foundation for estimating the applicant's profitability.

Predicting profitability requires estimates of both the probability of default and the likely usage of the store card. Whereas modelling default is the traditional task of credit scoring, estimation of usage is far less common. Time to the second purchase is considered as the measure of the card usage and dependencies in application and behavioural data are examined that can be used for predicting the customer's future behaviour.

Generic models are found to perform well across three countries under different modelling approaches and in different applications. In predicting default they are competitive with the national models, whilst in other applications generic models demonstrate marginally inferior results. However, harmonisation of the data available for the analysis is likely to further enhance the predictive power of generic models and expand the possible scope of their application.

Acknowledgements

I am indebted to my supervisors Professor Jonathan N. Crook and Dr Jake Ansell for their guidance and advice in preparation of this thesis.

I am very grateful to The University of Edinburgh and Credit Research Centre for funding this research. I am deeply obliged to the company that provided the data for this analysis for all the assistance with data preparation.

Many thanks to Maria Stepanova for setting a wonderful example and for encouragement.

I can never thank my family enough, especially my daughter, for patience and support.

Finally, I would like to express my most sincere gratitude to all my teachers, both in Russia and the UK.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AFT – accelerated failure time

AUROC - area under the Receiver Operating Characteristics curve

CRA - Community Reinvestment Act

CRA - credit reference agency

DA - Discriminant Analysis

DDA - Disability Discrimination Act

EC - European Community

ECOA - Equal Credit Opportunity Act

EU - European Union

FCRA - Fair Credit Reporting Act

FICO - Fair Isaac Credit Score

GA - Genetic Algorithm

GLBA - Gramm-Leach Bliley Act

HMDA - Home Mortgage Disclosure Act

HRA - Human Rights Act

KM – Kaplan-Meier

k-NN - k Nearest Neighbours

Lin Reg - Linear Regression

LP - Linear Programming

LR - Logistic Regression

LRB – logistic regression behavioural (model)

MS - Member States

NN - Neural Networks

OCC - Office of the Comptroller of the Currency

OFT - Office of Fair Trading

PH – proportional hazards

PHAB- proportional hazards behavioural (model)

RPA - Recursive Partitioning Algorithm

RRA -Race Relations Act

SDA - Sex Discrimination Act

SEA - Single European Act

SVM - Support Vector Machines

WOE - weights of evidence

Chapter 1. Introduction

1.1 Preface

With 12 countries joining the Euro, the exchange rate risk faced by credit applicants has been removed for those who earn and spend within these 12 Member States. This offers potentially an extremely attractive possibility of market expansion for credit institutions, provided they can manage the risk related not only to the residents of their own country, but also to applicants from other Euro States.

Credit scoring is the collection of techniques used for risk assessment in consumer credit. Its aim is to construct a classification rule that distinguishes between 'good' and 'bad' credit risks according to some specified definition. The accuracy of classification may be affected when the overall population comprises heterogeneous subgroups. It is not uncommon in credit scoring to develop separate models for each of such subgroups.

This thesis addresses several issues that arise in anticipation of the single European market in financial services. First, how different are the national credit risk patterns across Europe? Second, how accurate will the classification be if residents of several European countries are scored with one model? And third, what is the magnitude of improvement in classification from segmentation, i.e. from building individual models for different nations.

In a wider context these questions translate into the comparison of performance between generic models and customised ones. Generic models are models developed to fit several geographically or socio-economically different populations. In contrast to this, customised models are developed on, and applied to, only one population.

The examination of classification accuracy of different models is complicated by a problem of data comparability across countries and by diverse legislative environments that credit scoring models operate in. This calls for an investigation of the impact of the EU harmonisation on credit scoring practice.

The last but not the least aspect of the research project is an examination of the differential effects the generic/customised models may have on the inclusion/exclusion of certain groups of credit applicants.

1.2 Credit scoring

Credit scoring provides analytical support for decision making in consumer credit and allows for automation of the loan granting process. By relating the observable characteristics of new applicants for credit to the known performance of the previous borrowers, a credit scoring system ranks the applicants according to their attractiveness to a lender. The attractiveness, which is commonly referred to as creditworthiness, is most often based on the probability that the potential customer repays the loan on time.

The philosophy of credit scoring is predictive, not explanatory. It can be viewed as the behaviourist tradition that seeks to establish relations between human actions and some manifestations associated with them, not necessarily the causes of these actions. The cause is hard to establish and very often is not possible to observe at the point of application for credit, therefore the approach of credit scoring is extremely pragmatic: the exact cause of non-payment or default may be unknown, but prediction is still possible based on what is known, namely, the past experience and the available information about the potential customers. Such information is supplied by the applicant (from the application form), by credit reference agencies (CRA), and in the cases of existing customers can be derived from the lender's internal records.

There is a distinction between application scoring, that is concerned with the decisions whether to accept the applicant or not, and behavioural scoring, that supports decisions related to existing customers and uses the information about their transactions or behaviour. For a more detailed overview of application and behavioural scoring see Crook (1997); Hand (1998); Lewis (1992a); Thomas (1998); Thomas (2000); Thomas et al. (2002).

The credit scoring problem can be viewed as a supervised classification problem, where the task is to find some function or model that is the best separator between certain pre-defined classes (see Hand (1997)). Most frequently the model is a weighted sum of the applicant's observed characteristics (age, marital status, etc.), data on the incidence of defaults in the area where the applicant lives, and his/her previous performance pattern that produces a score. The score is a summary of the applicant's creditworthiness and reflects his or her ranking relative to other applicants. The classification into 'good' and 'bad' is achieved by comparing the score to a predetermined threshold or a cut-off level.

The credit scoring problem can be solved by a number of approaches. These include classical discriminant analysis, regression analysis, decision trees, neural networks, genetic algorithms and support vector machines. Each approach has its advantages and disadvantages, and their comprehensive discussion is given in Thomas (1998), Thomas et al. (2002), Henley and Hand (1996), Boyle et al. (1992), Davis et al. (1992), Yobas et al. (1997), Srinivasan and Kim (1987), Desai et al. (1997), Baesens (2003) and Schebesch and Stecking (2003). However, Table 1.1. shows that classification accuracy of different approaches is very close.

The numbers in Table 1.1 present the percentage of correctly classified accounts with the same acceptance rate, and should be compared across the rows only, since the authors used different definitions of 'good' and different cut-off levels.

Table 1.1 Comparison of classification accuracy for different scoring approaches (Thomas (2000), Baesens (2003)).

Authors	Lin Reg/ DA	Log Reg	RPA	LP	NN	k-NN	SVM	GA
Henley (1995)	43.4	43.3	43.8	-	-			-
Boyle et al. (1992)	77.5	-	75	74.7	-			-
Srinivasan and Kim (1987)	87.5	89.3	93.2	86.1	-			-
Yobas et al. (1997)	68.4	-	62.3	-	62.4			64.5
Desai et al. (1997)	66.5	67.3	67.3	-	64.0			-
Baesens (2003)[1]	74.4	74.4	74.8	74.8	75.0	74.8	74.8	-

[1] This study reports a range of results that vary depending on the type of the classifier used, on the dataset it was developed and tested on, and on the cut-off chosen. Table 1.1 gives results for the dataset UK1 assuming a cut-off of 0.5. Where several classifiers of the same type were used, the best predicting classifier is chosen.

So the choice of algorithm will depend on some additional features that any given approach can offer, its suitability for the problem under consideration and for the dataset used for modelling.

Linear programming is good for handling any constraints or conditions that may be desirable in a scorecard, e.g. to give older people equal or greater scores than scores of other age groups - the requirement that follows from credit regulations in the USA. Classification trees are effective in modelling the interactions between characteristics, whereas for regression models interactions have to specified in advance. Neural nets and support vector machines are particularly well suited for problems with non-linear relations. The nearest neighbour approach provides the possibility of developing models that are continuously updated and this allows one to adapt the scorecard to changing populations. Genetic algorithms screen a wide range of different models in search of the optimal one.

Regression approaches are among the most popular ones, since they offer the possibility of selecting statistically significant characteristics and performing collinearity tests. Logistic regression is favoured by the industry and has been chosen as the method of modelling in Chapter 4 of this thesis. Logistic regression overcomes the limitations of linear regression where the dependent variable can take values from $-\infty$ to $+\infty$, whereas the observed probability can range only between 0 and 1. It also gives less weighting to extreme values, and therefore is less sensitive to them than linear discriminant analysis or linear regression.

However, although one may intuitively expect logistic regression to give a greater classification accuracy than linear regression, Henley (1995) showed this is not the case. The reason lies in the fact that the central part of the logistic function is linear, and non-linearity starts in the regions where it is not difficult to separate goods from bads (Thomas (2000)), given that reasonable discrimination is possible.

One of the problems inherent to credit scoring is reject inference, which arises from the fact that the performance of rejected applicants is not known. There is a belief that developing the model only on the subset of accepted applicants may lead to a bias in estimation. There are several methods for incorporating rejected cases into the model, these include augmentation, parcelling, bivariate probit, multiple imputation and other methods (Hsia (1978), Joanes (1993)).

These methods have been criticised by Hand and Henley (1993), and the most recent papers by Banasik et al. (2001a), Banasik and Crook (2002) and Banasik and Crook (2003) showed that reject inference gives a very modest improvement, or even no improvement at all. Hand and Henley (1993) also argued that direct models of posterior probabilities (logistic regression among them) will give unbiased parameter estimates even if developed only on accepted cases, provided all the variables used in the previous accept-reject model are included into the new model to be estimated. Following these results, reject inference is not pursued in this thesis.

Traditionally credit scoring has been concerned with relating the applicant's observed characteristics to the probability of default within some specified period of time. However, it has been shown that the default probability is not necessarily a good indicator of the profit the applicant will generate (Hopper and Lewis (1992), Leonard (1997)). Some high risk applicants can generate a significant profit if they use the credit product actively and pay interest and charges for long enough before defaulting. On the contrary, low risk applicants may pay the full balance every month, thus keeping the revenues from such 'good' accounts low.

So recently the focus of credit scoring has been shifting from default to profit. In terms of the profitability of an account, one of the most important aspects is its lifetime, i.e. time before the borrower defaults or closes the account. Modelling the lifetime is the domain of survival analysis. Whilst survival analysis is a well-established technique in medical research and reliability, there have been only a few applications in credit scoring (Narain (1992), Banasik et al. (1999), Stepanova and Thomas (2001), Hand and Kelly (2001), Till (2001), Baesens (2003)). These studies found that in terms of classification accuracy the survival analysis approach is competitive and in certain applications superior to logistic regression. The advantage is that survival analysis provides an indication of both a customer's risk level and profit. This can be used to aid decision-making in relation to any individual borrower and also for debt provisioning based on the estimates of default levels over time.

Previous analyses investigated mainly fixed-term credit products. This thesis extends the application of survival analysis to the area of revolving credit (Chapters 5 and 6) and tests the performance of generic models under a number of modelling approaches.

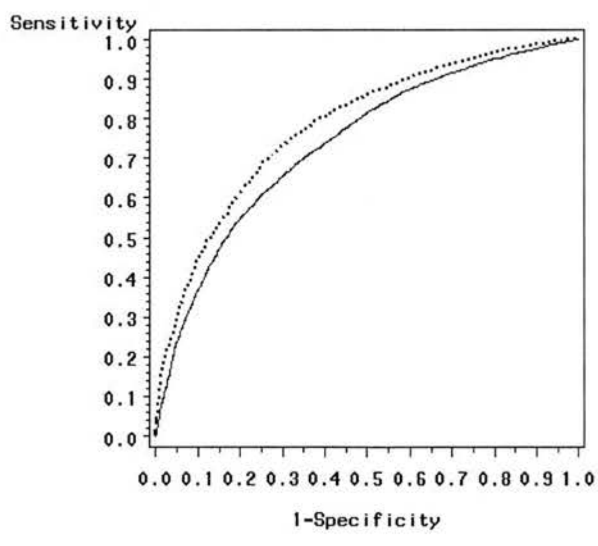
1.3 Measuring classification accuracy

All datasets used in the analysis were randomly split into training sets (70%) on which the models were developed, and hold-out sets (30%), reserved for testing the quality of prediction. It is well known that testing the classification accuracy of a rule on a sample it was trained on gives a biased indication of its performance.

The performance of models was compared using two measures: area under the Receiver Operating Characteristics curve (AUROC) and percentage of incorrectly classified accounts.

The Receiver Operating Characteristics (ROC) curve originated from the Theory of Signal Detection and initially was applied to quantify the ability of a system to detect a signal from noise when the two-alternative forced choice technique was used. The two alternatives were 'signal plus noise' and 'noise', and the distinction between the two classes could vary arbitrarily depending on how much noise was allowed to intervene with the signal. Sensitivity (or the true-positive fraction) is plotted on the *y*-axis against 1-Specificity (or the false-positive fraction), and this is done for all definitions of a 'signal'. Sensitivity is (the number of successful detections of a 'signal' according to a given definition)/(the number of 'signals' in total). Specificity is (the number of failures to classify 'noise' correctly)/ (the number of 'noise' instances in total).

Figure 1.1 Example of ROC-curves



Subsequently this method was used to judge the discrimination ability of various statistical methods in application to a number of different classification problems. In the context of credit scoring it is a plot of the proportion of correctly classified good accounts (P_{CG}) against the proportion of incorrectly classified bad accounts (P_{IB}) for all levels of predicted probability, or in other words for all possible cut-off levels *(c)*. Therefore it provides a measure of classification accuracy which is not dependent on any threshold or acceptance rate.

The perfectly discriminating system would have a graph joining (0,0) to (0,1) to (1,1) because

(P_{CG}) = [0;1] and (P_{IB}) = 0 for $c \leq P_G$

(P_{CG}) = 1 and (P_{IB}) = (0;1] for $c > P_G$, where P_G is the proportion of actual good accounts in the sample. A system giving no discrimination would correspond to a diagonal line. The dotted curve in Figure 1.1 corresponds to a system with better classification performance than the solid line.

The area under the curve offers a convenient measure for comparison. Alternatively, one may use the Gini coefficient, which can be easily obtained once the area under the curve is known (Hand (1997)):

$G = (A\text{-}1/2)*2$,

where A is AUROC.

It was shown by Bamber (1975) and Hanley and McNeil (1982) that conceptually AUROC corresponds to the Wilcoxon or Mann-Whitney or U statistic. This statistic estimates the probability, θ, that a ranking of a randomly selected bad account will be less than or equal to a ranking of a randomly selected good account.

This property allows for testing the significance of the difference between 2 or more ROC-curves. DeLong et al. (1998) extended the methodology to correlated ROC-curves, i.e. to curves that were generated by different tests or classification rules applied to the same sample. Using a method of structural components (Sen (1960)) the variance-covariance matrix of a vector of U-statistics is re-estimated. This is equivalent to jackknifing, because re-estimation is done by pairwise deletion of 'goods' and 'bads' from the sample.

DeLong et al. (1998) provided the formulation of the test of significance that can be applied to a vector of areas under correlated ROC-curves:

$$(\hat{\theta}-\theta)\boldsymbol{L}'[\boldsymbol{LSL}']^{-1}\boldsymbol{L}(\hat{\theta}-\theta)',$$

where $\hat{\theta}$ = a vector of statistics representing the areas under ROC-curves,

S = its estimated covariance matrix,

L = a contrast (a row vector of coefficients showing the sequence of comparison of elements of $\hat{\theta}$).

The test has a chi-square distribution with degrees of freedom corresponding to the rank of LSL'.

The second measure used for comparison is based on the confusion matrix, which presents the counts of good and bad accounts correctly and incorrectly classified by the model.

Table 1.2 Example of confusion matrix

		Actual Class	
		Good	Bad
Predicted	Good	a	b
Class	Bad	c	d

The use of the matrix requires the choice of a cut-off level or acceptance rate. For the purpose of this analysis the cut-off level was fixed at the level of the default rate in the hold-out sample, that is, such that the observed proportion of bads equalled the predicted proportion of bads in the hold-out sample.

The sum of percentages for incorrectly classified goods and incorrectly classified bads (error rate) was chosen as the measure for comparison:

$$\text{Error rate} = \frac{\boldsymbol{c}+\boldsymbol{b}}{\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}+\boldsymbol{d}}$$

Hand (2003) argued that the most relevant measure for assessing the scorecard performance is the 'bad' rate among the applicants predicted as 'good' by the model and hence accepted. The difference between bad rate and error rate is that the former measures one type of error, whereas the latter takes account of both types of error. Which measure is preferable depends on the type of question that is being asked.

1.3 Thesis overview

This thesis is structured in the following way. Chapter 2 looks at the effects of the anti-discrimination and data protection legislation in application to credit scoring. The existing regulations in the EU and USA together with legal and economic concepts of discrimination are reviewed. The analysis of previous research gives grounds to conclude that the existing legal restrictions on information, while impairing the quality of risk assessment, has failed to provide any tangible benefits to protected groups.

Chapter 3 presents a review of the previous research on generic models/ segmentation in credit scoring. The majority of studies demonstrate the superior prediction of customised/ segmented models over generic ones.

Chapter 4 analyses the differences in credit risk patterns of Belgium, Germany and the Netherlands and presents the logistic regression generic model that produces a good prediction for all three countries, comparable to that of national models. However, the applicants accepted by generic and customised models are not the same, and the implications for credit applicants from using a particular model are investigated.

Chapter 5 addresses the possibility of incorporating the time perspective into the analysis of credit risk and the differences in national patterns of default over time. Survival analysis methods (accelerated failure time and proportional hazards models) are compared to logistic regression in predicting the probability of default, and generic models are compared to the national ones.

Chapter 6 explores the possibility of predicting the future usage of the card. Time to the second purchase is considered since the interest lies in the timings of customer spending behaviour that provide the information on how quickly the lender's intervention is required in order to retain a profitable customer or to prevent the loss from an unprofitable one. The proportional hazards (PH) generic model is benchmarked against the proportional hazards national models.

Chapter 7 provides a summary of the findings and indicates future areas of research.

1.5 Contributions

This thesis addresses a crucial problem which is very likely to face European and eventually British banks in the near future, when households in countries of the European Monetary Union realize that the exchange rate risk they once faced has been removed, and start comparing the interest rates and loan terms across Europe rather than just within their own country. Institutions which fail to see this new market opportunity and to accurately assess the risk of lending to such households will yield market share to those who react quickly and adjust their credit-granting practice accordingly.

To maintain a competitive position and gain a competitive advantage banks must develop ways of assessing the creditworthiness of applicants from other countries. This presents a number of problems, such as the legal acceptability of predictor variables that can be used in such 'foreign' scorecards, including the issue of the very possibility of building separate models for each different EU country. In some countries the legislative context may make the generic models the only legally acceptable models. Therefore, it is important to know how the performance of generic models compares to specific national scorecards.

This thesis provides the first critical review of the differences in legislation relating to credit scoring models between the USA and European countries, in particular, the differences in laws on discrimination and data protection. It is argued that legitimate lending decisions are not necessarily 'fair'. Also the aspect of inclusion/exclusion has been considered in the use of generic versus customised models and observations have been made on the likely social impact.

Second, it provides a much more comprehensive analysis than previous literature of the comparative performance of generic European scorecards against national customised models. It shows that national models using the same variables as a generic model are only minimally more accurate than the corresponding generic model. But national models that use the full range of data available for the relevant countries are superior to generic models. This highlights the necessity for harmonisation of data across the EU.

Third, the thesis presents the first cross-country comparison of the application of survival analysis to predict when a borrower defaults. It shows that survival analysis generic models are competitive with the logistic regression generic models. No previous research has made such a comparison.

The desirability of offering a store card to an applicant depends not just on the chance the borrower will default but also whether s/he will use it. The fourth contribution of this thesis is to present the first inter-European comparison of survival analysis to predict when a store card holder will make a second purchase. Such a comparison has not been addressed in previous literature.

This thesis uses a unique and up to date dataset, which is particularly suited to reveal inter-country differences because it relates to the same type of credit product. The analysis therefore identifies inter-country differences in repayment and purchasing behaviour accurately because it controls for the differences in repayment behaviour that arise from the type of product.

In general, the thesis provides insight into national differences in credit risk patterns, contributes to an understanding of the harmonisation aspects of credit scoring practice in an integrated Europe, and investigates the performance of generic models across several countries and across different applications.

1.6 Summary

This Chapter presented the problems that were investigated and gave an overview of the paradigm of the research area (credit scoring), within which the problems were addressed. We can now proceed to discussion of the process of integrating the European financial markets into one entity, the difficulties associated with this process and its likely implications for credit scoring.

Chapter 2. The legislative environment of credit scoring models

2.1 Introduction

The personal character of the information used in automated decision-making gives rise to a number of legal and ethical questions as to what is appropriate to use in a credit scoring model (scorecard). The basis for resolving these questions is laid down by the relevant legislation that seeks to promote social justice policies and to ensure the confidential handling of personal information. Hence, it was pointed out by Hand (1998) that 'statistical scoring and classification systems all have to work within the boundaries defined by such legislation. The problems are thus not merely ones of mathematical optimisation, and neither are the solutions entirely clear cut.'

The problems of the legislative influence on credit scoring models become even more complicated when the lenders operate on an international level and need to consider several levels of regulations and to take into account the national differences in legislative frameworks. This is especially evident in Europe, where the process of economic and political integration has already brought to light some problems that have not received substantial attention before, such as whether the principal of non-discrimination on national grounds implies the prohibition to use nationality as a variable in credit scoring models.

This chapter addresses the impact of EU harmonisation on the practice of credit scoring. Given the complexity of the subject we wish to restrict ourselves to one aspect behind these differences – legislation. It seems to be justified for the following reasons:

1. The very process of the European integration is manifested in and implemented through certain legislative provisions.
2. Whilst the composition of scoring models is affected by a whole complex of different factors, and their impact is difficult to establish, the legislation has relatively direct and traceable effects, although sometimes (as it will be shown later) these effects may be complicated and controversial.

3. At the same time the legislature is influenced by the complex blend of the national political, economic and cultural factors and international interactions. To a certain extent, the national regulations are shaped by the same forces that manifest themselves in the diversity of the national credit risk patterns.

Against this background the objectives of this chapter are: first, to assess the nature of differences in national laws regulating consumer credit and the level of harmonisation achieved at the EU level. Second, we aim to establish whether the existing legal framework provides for effective credit risk assessment and smooth operation of the single European market in consumer credit.

Section 2 of the Chapter gives an overview of the legislation aimed at the creation of the single European market and the current state of the single market in financial services. It also investigates the national differences in the regulatory framework and their effects on the quality of credit risk assessment and functioning of the single market.

The legal measures can be broadly classified into two groups:

- Measures that follow from anti-discrimination legislation and seek to protect the principles of social justice;
- Measures that follow from the legislation protecting personal information.

Section 3 investigates the impact of the anti-discrimination law on credit scoring, Section 4 looks into the importance of CRA information, and outlines the national differences in data protection and credit referencing. Section 5 concludes.

The legislative impact on credit scoring has so far received very limited attention in the European context, but there have been a number of studies addressing this issue in the U.S. Thus, references to the US experience constitute the significant part of this Chapter, and this seems to be justified by the following reasons:

- The idea of the European integration was largely inspired by the example of the United States. The expression 'United States of Europe' goes back to 1814, when Henri Saint-Simon advanced the idea of the peace and unity through the integration (Urwin (1995)). Since then explicit and implicit references to the United States of America became common in the European context.

- The USA is also the country where credit scoring originated, and it is most advanced in many areas related to credit. So it is appropriate to use the USA as a 'benchmark' in the credit scoring context.
- In spite of some notable differences between the US and EU law, it will be shown that there are certain parallels in how both legislative systems address issues of discrimination and data protection. That allows for the conclusion that the US experience is relevant in the European context.

The Chapter concludes with some considerations of the adequacy of the existing legal framework for effective credit risk assessment.

2.2 The European Community and the Single Market

2.2.1 Principles of European Community law

The very idea of Pan-European scoring becomes relevant only in the context of an integrated market with a free flow of financial services across national borders. So first, it is necessary to review the current state of affairs in order to understand the basic features of the environment that a European scoring model would operate within.

The relationship between the European Community (EC) and its Member States (MS) is unique in the context of international law. EC law constitutes a completely new legal order, half way between international and domestic laws. It has limited the sovereign rights of the MS, but they still enjoy a high degree of freedom, and that is different from the legal order of federations or confederations. The difference from the international treaties lies in the fact that the EC laws affect not only MS but also their nationals.

The fundamental principles that regulate the complex relations between the EC and its MS are those of supremacy and subsidiarity. The principle of **supremacy** of EC law over national law implies that all national rules which conflict with the EC law must be disabled and all national courts have an obligation to ensure that this requirement is implemented in practice.

However, the principle of **subsidiarity** was introduced as a political compromise, a counter-measure against anticipated over-centralisation. It means that the EC and MS share the competence in certain areas, including such areas as economic and social cohesion and consumer protection.

Throughout this chapter it will be shown that there is no strict separation between the powers of the EC and MS, and the EC law is essentially the law of a compromise that tries to pursue integration and cohesion and at the same time to retain the uniqueness of each MS.

2.2.2 Single Market in financial services

One of the main ideas behind the establishment of the European Community was the creation of an integrated internal market. It is envisaged that the single market based on principles of free movement of goods, persons, capital and services will foster competition, thus consumers will benefit from wider choice and lower prices.

The year of 2005 represents the deadline for the integrated market in retail financial services as set by the Lisbon European Council.

The following legal measures represent the most important milestones on the way to the single market in consumer credit:

- Council Directive for the approximation of the laws concerning consumer credit (87/102/EEC, Council of the EU (1987) amended by Directive 90/88/EEC, Council of the EU (1990) and Directive 98/7/EC, European Parliament (1998)), which aimed to achieve a certain level of convergence of the rules related to lending and to ensure a high degree of consumer protection. The Directive laid down the requirement of official authorisation (licensing) for credit-grantors.
- Second Banking Co-ordination Directive (89/646/EEC, Council of the EU (1989)), which formulated the principle of the single licence that allows banks and other credit institutions to set up branches and offer services throughout the Community.
- Data Protection Directive (95/46/EC, European Parliament (1995)), which will be discussed in detail in Section 2.3.

- European Parliament and Council Cross-Border Credit Transfers Directive (97/5/EC, European Parliament (1997)), aimed at the establishment of the minimum information and performance requirements for cross-border credit transfers so as to ensure that funds can be transferred from one part of the Community to another rapidly, reliably and inexpensively.
- E-commerce Legal Framework Directive (2000/31/EC, European Parliament (2000)), aimed at ensuring the free flow of on-line services across the Community.

These Community measures are meant to enable the unobstructed flow of financial services across national borders and are based on three basic principles:

- essential harmonisation in all Member States of the laws and practices governing access to financial services,
- home-country control, reinforced through co-operation between national supervisory authorities,
- mutual recognition by national supervisory authorities of the rules and regulations in the countries of origin of the banks operating on their territory. It means that providers of financial services that comply with the regulations of the Member State of their registration, may operate in other MSs without any further restrictions.

These principles imply that a lender holding a license in one MS can set up branches and offer services throughout the Community, and the lender's activities will be regulated by the legislative framework of its home state. However, in practice this has not always worked for the following reasons.

First, the main legal instruments of harmonisation are Directives that formulate the final target that should be achieved, but it is up to MSs to decide how this target should be reached. This results in a certain divergence in transposition of Directives. The situation is further aggravated by delays in implementing the Community legislation by MSs. In May 2003 the Commission reported that 8.8% of all Internal Market Directives have not yet been transposed by all MS (European Commission (2003)).

Second, it is recognised that the objective of harmonisation, or creating 'the level playing field' for all traders throughout the EC, can result in a minority of states with stricter rules being forced to bring their regimes down to a harmonised standard. This made the Community adopt the 'minimum harmonisation' formula in many cases. Rather than setting a single Community rule as both floor and ceiling, the Community measure (Directive) acts as a floor.

An example of the application of this rationale is Directive 87/102/EEC which specifically addressed 'the approximation of the laws, regulations and administrative provisions of the Member States concerning consumer credit'. The Directive had a very limited effect on actual harmonisation, due to the fact that Article 15 (Minimum Clause) of this Directive gave the right to MSs to adopt more stringent legal rules for consumer protection. As a result, most Member States have systematically used this right and went beyond the provisions of the Directive. This created a paradox: the Directive was initially designed to ensure harmonisation, but achieved quite a modest harmonising impact.

Third, MSs have the right to derogate from the principle of mutual recognition on the grounds of 'general good' that include:

- the effectiveness of fiscal supervision,
- the protection of public health,
- the fairness of commercial transactions,
- the defence of the consumer.

This means that lenders still have to consider some regulations of their host MS.

To summarise, the single market in retail financial services is not a reality yet. In spite of numerous Directives, the level of harmonisation in consumer credit markets is modest, and lenders still have to take into account some differences in the regulations of the MSs they operate in.

The Commission has adopted the Financial Services Action Plan (European Commission (1999b)), and one of the tasks listed there, is the amendment of the Consumer Credit Directive (87/102/EEC). In September 2002 the proposal for a new Directive was presented by the Commission, but it was not adopted yet (European Commission (2002)).

The new Directive was designed in recognition of the fact that diverse national rules reduce cross-border transactions. It offers a comprehensive set of common regulations and explicitly prevents MSs from adding new rules. The proposal incorporates the concept of 'responsible lending' and places an obligation on lenders to conscientiously assess the borrower's ability to repay the loan. So more changes are anticipated in the very near future.

However, the proposal does not contain clarification on what information can be used for credit risk assessment, so the information available for model building in credit scoring is subject to restrictions following from the general provisions of the anti-discrimination and data protection legislation. Although there are Community measures in this field, it will be shown that national divergence is still quite significant.

2.3 Anti-Discrimination Law and its affects on credit granting decisions

2.3.1 Legal and economic definitions of discrimination

The Oxford English Dictionary gives the following definition of discrimination:

"Discriminate - a) to make a distinction; to perceive or note the difference; to exercise discernment.

b) to discriminate against; to make an adverse distinction with regard to; to distinguish unfavourably from others." (The Oxford English Dictionary (1933))

Clearly there are two different meanings of this word[2] and this is reflected in the economic literature (Yinger (1997)). These two meanings are:

- statistical (or economic) discrimination, which is objective;
- taste-based (or non-economic) discrimination, which is subjective.

[2] It should be noted that two different meanings of the word 'discrimination' do not exist in all languages; 'discrimination' was borrowed from English by other languages bearing the negative connotation only, its neutral meaning was dropped. So when 'discrimination' is used on the international level, it implies 'adverse or unfair treatment.'

It can be argued that any form of selection involves discrimination in one or both of the senses above. Whether it is justified or not, can be decided on the basis of reasons used to select these particular applicants and to reject the others. Where human judgement is involved, discrimination arises from preferences or prejudices. According to taste-based discrimination theory (Becker (1971); Peterson and Peterson (1978)) the person who discriminates must pay extra in order to have the privilege of not dealing with certain groups of people, but there will be a point beyond which it becomes too expensive to discriminate.

This theory does not apply to credit scoring since there is no personal judgement involved. But statistical discrimination (Avery (1981); Phelps (1972)) which arises from the lack of information necessary to calculate the degree of risk, does apply to credit scoring. The estimate of risk of a particular individual is based on the risk estimates of groups that this individual belongs to. Therefore the decision is based upon the assumption that people with similar characteristics will behave in a similar way. This is objective discrimination, since it is based on past experience. So if in the past people of a certain gender or race proved to be poorer credit risks, this gives grounds to believe that there is a higher probability of these persons not paying back on time.

The lender might reasonably argue that this type of discrimination is economically rational and hence market forces will not eliminate the discrimination. It would constitute discrimination if members of a demographic group are more likely to be rejected by the lender than those from other groups with similar characteristics besides group membership (Yinger (1997)). Hence the membership of the group is influencing the default probability.

Avery (1981) differentiates statistical discrimination into two types: the 'endowment' effect and the 'mean shift' effect. The endowment effect results from a difference in the economic variables between the groups and hence it can be argued that it does not constitute discrimination but a response from these differences. The mean shift effect describes those cases where group membership provides information about default beyond that supplied by the economic variables. Practically, though, it is not always possible to differentiate between these two effects since group membership is often correlated with economic measures.

In contrast to the subjective/objective duality of economic theory, the legal point of view is based on the principle of equal treatment and tries to identify whether the selection process breaches this principle. Or in other words, would the selection results be different e.g. if the person was of a different race other things being equal.

The breach of the principle of equal treatment becomes illegal only when 'equal treatment' or 'non-discrimination' is linked to certain grounds, e.g. gender, race, ethnic origin, disability, and when the certain areas of application are specified.

> 'For an action to be discriminatory in law, there must first be a law, and this must define the prohibited grounds of action, the persons protected by the law, and the circumstances in which they are protected' (Banton (1994)).

The law also acknowledges the existence of different types of discrimination, but its focus is centred on the distinction between 'purpose and effect'. So one type of discrimination (most commonly referred to as direct discrimination) arises when the selection process intentionally separates one group from the other. The second type (indirect discrimination) arises when there is no apparent intentional separation, but the results of the selection process disadvantage certain groups.

As we shall see views on what constitutes discrimination are influenced by cultural traditions, so the perceptions of inadmissible grounds for discrimination can vary from country to country which results in different interpretations of the principle of equal treatment and what is justifiable.

2.3.2 EC anti-discrimination legislation

It was pointed out earlier that the non-discrimination law should specify the grounds for equal treatment and areas of application. The grounds are derived from the human rights enshrined in national constitutions and international treaties. There are two stereotypical ways of formulating human rights: as an obligation for a contracting party or as a right for an individual. The former involves the obligations to create certain conditions or to pursue certain policies (sometimes they are called 'programmatic rights'), but individuals cannot rely on them before a national or international court.

The second category creates so-called 'legal' or 'justiciable' rights. These are formulated in such a way that they can be invoked in court by individuals ('self-executing' effect). One of the main conditions of self-executing treaty provisions is that they be clear and unconditional and require no further government action. Traditionally, civil and political rights were seen as justiciable, whereas economic and social rights were generally regarded as 'programmatic'. (Banton (1994) refers to the latter as 'the view that human rights begin after breakfast' and the former as 'the view that human rights begin in the police station').

However, there is no longer a precise distinction between the two categories, and the European Community is (or was) mainly an economic organisation, which makes the protection of social rights at least as relevant as the protection of civil rights.

It should be noted that there is no such a thing as 'a right to credit'. Nevertheless, the Council of the European Union acknowledges that 'credit is a driving force of economic growth and the welfare of consumers'(Council of the EU (2001)). It becomes subject to anti-discriminatory regulations, when the law includes 'access to services and goods' in the scope of application.

Initially the EC law directly addressed two issues:

- non-discrimination on the grounds of nationality;
- equal treatment of men and women.

The principle of non-discrimination is one of the fundamental principles of EC Law. Article 12 of the Treaty establishing the European Community (ex. Article 6), as modified by the Amsterdam Treaty, explicitly prohibits discrimination on the grounds of nationality: 'Within the scope of application of this Treaty, and without prejudice to any special provisions contained therein, any discrimination on grounds of nationality shall be prohibited.' (Treaty (2002)).

Since this principle is crucial for the functioning of the single market, it has been given the strongest power possible: Article 12 is directly effective (i.e. no further legal acts are required to enforce it) and it has both vertical and horizontal effect (i.e. it is binding for the public authorities, private entities and individuals). It should be noted, though, that Article 12 does not automatically create the prohibition to use 'nationality' in credit scoring. It gives the right to the national authorities to take measures to eliminate national discrimination (if they believe that it is occurring) and it also gives the right to an individual to rely on this Article in Court if s/he believes that s/he was treated unjustly. So interpretation by a Court or national authorities is important.

The EC Law also proclaimed the equality of men and women: '...the Community shall aim to eliminate inequalities, and to promote equality, between men and women.' (Article 3 EC, Treaty (2002)).

However, apart from nationality and sex, the EC Law originally did not cover equality on any other grounds, neither did it cover any other human rights, although the European Court of Justice (ECJ) developed the formula that the protection of human rights is enshrined in general principles of the EC Law[2].

The situation changed after the emergence of the political union - the European Union. The Amsterdam Treaty (Treaty (1997)) added Article 13 which provides for a platform for expanded anti-discrimination policies. Its text reads: 'the Council ... **may** take appropriate action to combat discrimination based on **sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation**[3]'. However, unlike the principle of non-discrimination on the grounds of nationality, Article 13 does not have direct effect, so it has to be implemented by the specific legal measures.

In legislative aspect Article 13 translated into two Directives:

- Directive 2000/43/EC of June 2000 implementing the principle of equal treatment between persons irrespective of racial and ethnic origin. The Directive goes beyond the labour market, covering the 'access to and supply of goods and services which are available to public' and distinguishes between 'direct' and 'indirect' discrimination. It has adopted the following definitions:

  'Concept of discrimination

  1. For the purposes of this Directive, the principle of equal treatment shall mean that there shall be no direct or indirect discrimination based on racial or ethnic origin.

---

[2] General principles of law, which do not necessarily have to be written down, function as a standard of good behaviour of, usually public authorities or governments, so that the human rights of individuals are not endangered (Usher (1998)).

[3] Bold added.

2. For the purposes of paragraph 1:

(a) direct discrimination shall be taken to occur where one person is treated less favourably than another is, has been or would be treated on grounds of racial or ethnic origin;

(b) indirect discrimination shall be taken to occur where an apparently neutral provision, criterion or practice is liable to affect adversely a person or a group of persons of a particular racial or ethnic origin, unless that provision, criterion or practice is objectively justified by a legitimate aim which is unrelated to the racial or ethnic origin of a person or group of persons and the means of achieving that aim are appropriate and necessary'.(Council of the EU (2000a)).

- A Directive establishing a general framework for equal treatment in employment and occupation (2000/78/EC), which expands the equality principles to the grounds given above (Article 13) in addition to the earlier proclaimed (Articles 3 and 141) equality of men and women in labour markets (Council of the EU (2000b)).

In addition, in the near future one may expect more changes in the anti-discrimination field. On 7 December 2000 the EU Charter of Fundamental Rights was signed and proclaimed by the Presidents of the European Parliament, the Council and the Commission at the European Council meeting in Nice. The Charter is not legally binding yet, since work is under way to incorporate it into the Treaty establishing the Constitution for Europe, that will set out, in a single text for the first time in European history, the whole range of civil, political, economic and social rights of European citizens and all persons resident in the EU.

Of particular interest is Article II-21 'Non-discrimination', which states:

'Any discrimination based on any ground such as **sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation** shall be prohibited.' (Draft Treaty (2003)).

This statement does not have a legal power yet, so in terms of access to goods and services it is only discrimination on race and ethnic origin that is forbidden at the EU level (following the implementation of Race Directive (2000/43/EC) in July

2003). But in some MSs the anti-discriminatory regulations are more stringent or cover a wider scope than the EU legislature. As can be seen from Table 2.1, Ireland, the Netherlands and the UK already have legal acts that outlaw discrimination in access to services and goods on grounds wider than race and ethnic origin.

For example, the UK has specific acts addressing discrimination on the grounds of gender (Sex Discrimination Act, SDA (1975)), race (Race Relations Act, RRA (1976), disability (Disability Discrimination Act, DDA (1995)) and religion (Human Rights Act, HRA (1998)). And even the definition of 'racial grounds' have a broader scope than the EU one: the UK 1976 Race Relations Act covers the grounds of race, colour, nationality (including citizenship), ethnic or national origin. The Act declares unlawful 'segregating a person from other persons on racial grounds' or direct discrimination in other words. At the same time the Act states that even when any requirement is applied equally to all persons, but 'the proportion of persons of the same racial group as that other who can comply with it is considerably smaller than the proportion of persons not of that racial group who can comply with it', this also constitutes discrimination. So if some minority ethnic group has lower income, shorter terms of living at the same address or working in the same job, compared to other ethnic groups, and these variables are included into credit risk assessment, and therefore, members of a disadvantaged group are granted less credit, then it also constitutes discrimination, unless it can be shown that the requirement is 'justifiable'. Such a requirement applies to 'colour' and 'nationality' (RRA (1976)).

In terms of race or racial/national origin, the recent amendment of the said Act (Regulations (2003)) implementing the EU Race Directive, removes the condition for ethnic group to be 'considerably smaller'. It states that even if the requirement is applied equally to all persons, but members of a certain ethnic group would be at a disadvantage, then it is necessary to show that this requirement constitutes 'proportionate means of achieving a legitimate aim'.

It seems that a general problem of the anti-discrimination provisions in Europe is that they allow for several interpretations. For example, in the case of the UK RRA, there is no indication of what is 'considerably smaller'. Equally, there is no explanation of what provisions would be regarded as appropriate and legitimate in credit risk assessment.

Table 2.1 National provisions[4] against discrimination on grounds of race/ethnic origin, religion/ belief, disability[5] , age or sexual orientation

Member State	Laws or collective agreements	Grounds covered	Scope
Belgium	Collective agreement of 6 December 1983 concerning the recruitment and selection of workers.	Racial or ethnic origin, religion or belief and age.	Employment.
Denmark	Act 459 of 12 June 1996 on prohibition against discrimination in respect of employment and occupation etc.	Racial or ethnic origin, religion or belief and sexual orientation	Employment.
Germany	Civil Service codes and the Works Constitution Act (Betr.VG)	Racial or ethnic origin, religion or belief.	Employment
Greece	No anti-discrimination law.	-	-
Spain	The Workers Statute Act ("Estatuto de los Trabajadores")	Racial or ethnic origin, religion or belief and age.	Employment.
France	The Labour Code	Racial or ethnic origin, religion or belief, disability and sexual orientation	Employment.
Ireland	Employment Equality Act 1998, Equal Status Act 2000.	Racial or ethnic origin, religion or belief, disability, sex and sexual orientation, age, marital status, family status	Employment, education, the provision of goods and services and the disposal of property and accommodation
Luxembourg	No anti-discrimination law.	-	-

[4] The table does not cover the constitutional provisions. All the MS have articles in their constitutions providing for the equality principles, except for the UK that does not have a written Constitution.

[5] Only legislation which makes it unlawful to discriminate against disabled people are mentioned in the table. Concerning the integration of the disabled people several Member States have systems of compulsory employment or quota schemes (Germany, Greece, Spain, Italy, Luxembourg, the Netherlands and Austria) while others rely on subsidies to employers.

Member State	Laws or collective agreements	Grounds covered	Scope
Italy	Law no. 300 of May 20, 1970 (Workers Statute) and Law no. 40 of 6 March 1998.	Racial or ethnic origin and religion or belief.	Employment (Workers Statute) and the provision of services and goods (Law no 40).
The Netherlands	The Equal Treatment Act(1994)	Racial or ethnic origin, religion or belief, sex, sexual orientation, age, disability, social status	Employment, advice regarding choice of education or career and provision of goods and services.
Austria	The Introductory Act to the Administrative Procedures Code (*Verwaltungsverfahrengesetze*).	Racial or ethnic origin and religion or belief.	Provision of public service and admission to public places
Portugal	No anti-discrimination law.	-	-
Finland	Act on Contracts of Employment, Act on Equality between Women and Men	Racial or ethnic origin, religion or belief, age and sexual orientation	Employment, healthcare
Sweden	Employment Protection Act (1982), Ethnic Discrimination Act (1999), Act on Discrimination of people with disabilities (1999), Act on discrimination on grounds of sexual orientation (1999), Act 2003:307	Racial or ethnic origin, religion or belief, handicap, sex and sexual orientation.	Employment, education, goods, services, housing
The United Kingdom	Sex Discrimination Act, Race Relations Act, Disability Discrimination Act, Human Rights Act	Racial or ethnic origin, religion or belief, sex, disability.	Employment, training, education, provision of goods, facilities and services, or management and disposal of premises.

Source: 'Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions on certain Community measures to combat discrimination' COM 564 final (European Commission (1999a)) and DG Employment and Social Affairs (http://europa.eu.int/comm/employment_social/fundamental_rights/legis/mslegInracequal_en.htm)

This problem is partially resolved by the *Guide to Credit Scoring* (OFT (2000)) compiled by the credit industry and approved by The Office of Fair Trading and Department of Trade and Industry. Although this document lays down the principles of appropriate standards for scorecard development and validation, it is not legally binding. Section 2.4 of the Guide states that 'Credit Scoring will not discriminate on the grounds of sex, race, religion, disability or colour.' It is not clear whether the statement covers direct or indirect discrimination.

So the final decision about whether any particular action constitutes discrimination, and what kind of discrimination, is left to the regulating authorities and the Court. We are not aware of any Court cases on discrimination in credit scoring in the UK.

However, there was a case in France, where the national data protection authority - Commission National de l'Informatique et des Libertes (CNIL) – has explicitly prohibited the use of nationality in credit scoring models in 1998 (CNIL (1998)). But subsequently this decision was overruled by the higher authority (Le Conseil d'Etat (2001))[6].

In France the final decision was achieved by balancing two objectives: the necessity to achieve an accurate risk assessment and the necessity to protect human rights. It is not straightforward which objective should be given a proirity. In this context, questions about how the anti-discrimination measures affect borrowers and lenders, and whether these measures actually achieve their intended purpose become of fundamental importance. These questions have already been addressed in the USA, where a considerable empirical evidence on the effects of the anti-discrimination law has been accumulated. The next section will outline the U.S. regulations.

[6] More details on this case are given in Section 2.4.2.3.

2.3.3 Anti-Discrimination Law of the USA

Unlike Europe, where restrictions on information in credit scoring have to be inferred from the general anti-discrimination regulations, the USA has a package of legal acts that deal specifically with the area of credit (Table 2.2). This can be accounted for by the large volumes of credit and the growing dependence of US consumers on it, which resulted in the situation where the public would like to view the access to credit as a right. (Lewis (1992b))

The most important US legal acts in the area of credit are outlined below, except for the Fair Credit Reporting Act (FCRA, 1970) that will be discussed in detail in Section 2.4.3.

The Equal Credit Opportunity Act (ECOA (1976)) and its implementing regulation (Regulation B) prohibit lending discrimination on the basis of **race, colour, national origin, age, gender, marital status, religion, receipt of public assistance, or exercise of rights granted by consumer protection statutes.** It distinguishes between judgmental and statistical scoring systems, and allows the use of age in the latter as long as its use does not disadvantage applicants over 62 years.

The Home Mortgage Disclosure Act (HMDA (1975) requires financial institutions to publicly disclose the number and dollar volume of home mortgage loans they make in metropolitan areas. Amendments in 1989 put an obligation on the lenders to collect and report data regarding race, gender and income characteristics of borrowers and applicants for mortgage loans. Previously the law required disclosure only of loans made, not applications. With the provision to disclose data not only on loans granted but also on loans denied, the Act became a tool to aid enforcement of the fair-lending laws. The section below on the Boston Federal Reserve Bank Study (Section 2.3.3.2) gives an example of how it has been implemented in practice.

The Community Reinvestment Act (CRA (1977)) encourages financial institutions to help to meet the credit needs of the communities in which they operate, including low- and moderate-income neighbourhoods, consistent with safe and sound banking operations. The Act requires that each financial institution's record in helping meet the credit needs of its entire community be evaluated periodically. That record is taken into account in considering an institution's application for deposit facilities, including mergers and acquisitions.

Table 2.2 Differences in legislation in the USA, EU and UK in application to credit scoring

Characteristics that are legally restricted	**USA**	**EU**	**UK**
Race/ colour/ ethnic origin	Equal Credit Opportunity Act (ECOA)	Data Protection Directive (95/46/EC), Directive 2000/43/EC on racial and ethnic origin equality/	Race Relations Act
Nationality/country of residence		Article 12 EC, Data Protection Directive (95/46/EC)	
Religion /belief	ECOA	Data Protection Directive (95/46/EC)	Human Rights Act
Health or sex life		Data Protection Directive (95/46/EC)	
Gender	ECOA	Proposal for a new gender equality Directive	Sex Discrimination Act
Age	ECOA		
Marital Status	ECOA		
Politics/ trade union membership		Data Protection Directive (95/46/EC)	
Disability			Disability Discrimination Act
Third party data	Fair Credit Reporting Act	Data Protection Directive (95/46/EC)	Data Protection Act 1998
Neighbourhood characteristics	Community Reinvestment Act, Home Mortgage Disclosure Act		OFT Report OFT (1992)
Public assistance income	ECOA		
Exercise of credit consumer rights	ECOA		

2.3.3.1 Overt discrimination, disparate treatment and disparate impact

In the US regulations the distinction is made between

- overt discrimination – explicit use of forbidden variables in scoring models,
- disparate treatment – judgmental or subjective discrimination, which may occur, when the score derived from the statistical model is judgementally adjusted,
- and the disparate impact (Office of Comptroller of Currency (1997)).

The joint agency Policy Statement on Discrimination in Lending (Office of Comptroller of Currency (1994)) explains what constitutes disparate impact. Even if a variable (characteristic) is not explicitly banned, if it leads to excessive rejection of borrowers of a certain race or gender or with respect to some other prohibited characteristic, the lender needs to show there is a 'business necessity' for using the variable and there is no equally effective way of making the credit decision.

So in principle there are parallels between the US and European definitions: overt discrimination and disparate treatment in the US correspond to direct discrimination in Europe, and disparate impact is similar to indirect discrimination.

There is also no clear definition of what constitutes 'business necessity'. The Bulletin 97-24 issued by the Office of the Comptroller of the Currency (Office of Comptroller of Currency (1997)) says that a variable used in a credit-scoring system will essentially be assumed to meet the business necessity test if it is statistically related to loan performance and has an 'understandable relationship' to the applicant's creditworthiness.

There is still a lot of subjectivity involved in interpreting whether a relationship is 'understandable' or not, and what level constitutes the right balance between statistical association and disparate impact.

Nevertheless, Mester (1997) believes that it seems to be generally accepted that a credit scoring model makes it easier than a judgmental decision-making for a lender to document the business reason for using a variable that might have a disproportionately negative effect on certain groups of applicants protected by law from discrimination. The significance of each weight in the model gives a measure of the marginal impact of each variable on the likely credit performance (given the other variables contained in the model). Also, a well-built model will include all allowable factors that produce the most accurate prediction of credit performance, so a lender using such a model might be able to argue that a similarly effective alternative was not available.

Therefore, for some time US lenders believed that credit scoring gave them the chance to fully comply with fair-lending legislature by removing subjectivity from the credit granting process. But the Boston Federal Reserve Bank study (Munnel et al. (1992); Munnel et al. (1996)) changed this complacency.

2.3.3.2 The Boston Federal Reserve study

The study analysed the data coming from loan applications for mortgages in the Boston area. The sample included all applications submitted by blacks and Hispanics and a random sample of white applicants. The characteristics were taken from the data collected under the HMDA and included each applicant's race, gender, income, and whether the application was accepted or denied. The denial rates showed substantially higher denial rates for black and Hispanic applicants than for white applicants. Both ordinary least squares and binomial logit techniques were used to estimate the probability of being denied a mortgage. The 1992 study concluded that minority applicants in the Boston area were three times more likely to be denied a mortgage loan than whites, and both whites and minorities were more likely to be rejected when the property was located in a minority neighbourhood, as would be expected if discrimination and redlining were occurring.

A key methodological problem in estimating the importance of the race coefficient arises, as pointed out by Yinger (1997), because many characteristics of the applicants and neighbourhoods are correlated with race. In other words, minority applicants tend to have poorer credit qualifications. Yinger believes that this correlation is a problem because it implies that if some important characteristics are missing this can result in serious omitted-variable bias in the coefficients of race indicators. If income is negatively correlated with the probability of loan denial and with race as well, the coefficient of race can be higher, if income is omitted from the model. Because this coefficient is the estimate of discrimination and redlining, one may conclude that there is discrimination or redlining when in fact there is none.

Since the 1992 Boston Federal Reserve study was heavily criticised for this omitted-variable bias, Munnel et al. (1996) improved on their earlier study by including a more comprehensive list of applicant and property characteristics. These 38 additional variables were collected by surveying the lenders in Boston, and the authors argued that they included all possible factors that lenders used for predicting default.

The results of this study indicated that minority applicants, on average, did have less wealth, weaker credit histories, and higher loan-to-value ratios than white applicants, and these disadvantages did account for a large portion of the difference in denial rates. Including the additional information on applicant and property characteristics reduced the disparity between the minority and white denials from the originally reported 18 percentage points to just 8.2 percentage points. It also reduced a relative rejection ratio of 2.8 to 1 to a relative rejection ratio of roughly 1.8 to 1. The improved study did not find any evidence for redlining on the basis of the racial composition of the area.

Nevertheless, the conclusion was that there was still enough evidence to suggest that the discrimination was occurring, although, on the other hand, the results of the improved study could be interpreted as supporting the idea of the omitted-variable bias. If the original disparity is reduced by including the additional characteristics into the model, perhaps, it can be reduced further with the addition of further variables.

Lewis (1992b) states that one can never predict with 100% certainty if any particular person will default, and that credit scoring models are not causal models, but associative ones. Since it is not known what causes a good or bad performance, the attempts can be made only to find facts that are closely associated with it. So to claim that all possible factors were collected that influence the credit performance is, perhaps, too optimistic.

However, the problem of correlation pointed out by Yinger is not limited to the omitted-variable bias, since race is correlated not only with omitted variables but also with variables included into the model that is used to estimate the significance of discrimination.

It was shown by Bostic (1997) that the race indicator was significantly correlated with all economic variables included in Munnell's analysis, although the correlation was not large enough to suggest that multicollinearity should be a major problem in estimation. Nevertheless, Bostic demonstrated that the minority and white populations did differ demographically and in terms of economic and social status. When the interaction terms between race and some of the financial variables were included into the model, the race indicator lost its significance. Further analysis lead Bostic to conclude that the claim that discrimination was a general phenomenon was refuted, since the racial differences in probability of being accepted existed only for 'marginal' applicants, i.e. applicants with low economic indicators and poor credit history. However, even for this group there was no evidence to claim that this difference was occurring due to discrimination rather than due to economic factors.

Finally, it should be noted that the studies aimed at revealing discrimination are trying to mimic the decision-making process of lenders. In this case it is not enough just to assemble the information that is available for making a credit-granting decision. There are different modelling approaches and techniques that may result in completely different models being built on the same dataset. One cannot claim that race is included into the decision-making process, without taking this into account.

Although this study was heavily criticized by economists (Brimelow and Spencer (1993); Day and Liebowitz (1998); Harrison (1998); Liebowitz (1993); Zandi (1993); Altman et al. (1981)), the policy-makers and public opinion supported the conclusions of the Munnel report. The influence of the Munnel report can be demonstrated by the following statements of Lawrence Lindsay, a member of the Board of Governors of the Federal Reserve at that time:

> 'I dismiss the view of conservative economists and bankers that … there are no racially based problems in mortgage lending… The study (Munnel et al.) may be imperfect, but it remains a landmark study that sheds an important light onto the issue of potential discrimination in lending.' (Liebowitz (1993))

The Boston Fed Study highlighted the important problem: the prohibition of certain grounds from credit scoring models does not eliminate the discrimination on these grounds. And 'equal treatment' does not automatically create an equal world.

A further aspect of the problem is that even if credit grantors do not discriminate between minorities, when assessing creditworthiness, minorities may believe they do. Crook (1999) found that some minorities, in particular blacks, Hispanics and single females are more likely than other groups to be discouraged from applying for credit, although they are not more likely to default. So eliminating discrimination requires more than a ban on certain information. The next section will discuss this issue further.

2.3.4. Impact on credit granting

Although the anti-discriminatory legislature can vary significantly from country to country, it shares the same objective: to provide protection to disadvantaged groups. In the context of consumer credit, this means to increase their access to it. In the US environment Elliehausen and Durkin (1989) investigated whether the ECOA achieved this goal.

While analysing the effects of the anti-discrimination legislature it is important to answer the following questions:

1) whether the removal of the prohibited characteristics impacts the predictive ability of credit scoring models;
2) whether the protected groups are poorer credit risks and require the protection;
3) whether the prohibition of certain characteristics increases the acceptance rates for the protected groups.

As far as the first question is concerned, a study on the effect of limiting information in credit scoring models was carried out by Shinkel (1980). He developed eight discriminant models, from which seven models excluded variables prohibited by the ECOA, and one model contained the prohibited variables. Each model was used to classify applicants in a holdout sample. His results indicated that exclusion of prohibited variables reduced the number of good loans accepted (depending on the variable that was excluded the percent of reduction ranged from 0.3% to 2.3%) and increased the number of bad loans accepted (from 0 to 2.6%) with a reduction in profitability of 2% to 16%.

On the other hand, Elliehausen and Durkin (1989) refer to studies that demonstrated the opposite result. Thus, Nevin and Churchill, Jr. (1979) and Shay and Genderton (1979) concluded that there were no significant differences in predictive ability with and without gender, marital status and age included. But such results can be explained by the presence of variables strongly correlated with the excluded ones, so that the variables remaining in the model 'proxy' the prohibited ones.

In the European context Platts and Howe (1997) attempted to develop a single generic model for five EU countries that represented different regions in Europe. In this way the authors wanted to apply one model to the whole of EU. Their analysis showed that the predictive power of the model went down significantly when no distinction was made between nationality or country of residence.

As for the creditworthiness of the protected groups, the evidence on relationship between race and default is controversial. Avery (1982) found that black applicants appeared less likely than other applicants to pay off their accounts as scheduled, even after controlling for applicants' financial and credit characteristics. Similarly, Boyes et al. (1978); Martin and Hill (2000) showed that racial minorities were more likely than whites to default, other things equal. On the other hand, Van Order et al. (1993) showed the contrary evidence for some parts of the USA.

But women and older borrowers were found to be less risky than other borrowers (Altman et al. (1981); Avery (1982); Boyes et al. (1978); Chandler and Ewert (1976) and Durand (1941)). So one can argue on the basis of these studies that some protected classes are actually *more* creditworthy than some unprotected groups.

Further, there have been studies suggesting that the prohibition of certain characteristics failed to increase acceptance rates for protected groups. The analysis by Chandler and Ewert (1976) suggested that separate risk profiles for male and female applicants could identify credit risk more precisely than a model which ignored an applicant's gender or one which allowed for only limited differences in male and female risk profiles. And what is more important, the acceptance rates for females were higher when gender was included into the model as a variable and also in segmented male/female models as compared to the model that did not distinguish between sexes. So one can conclude that the ECOA appeared to disadvantage rather than benefit female applicants.

A more recent study by Banasik et al. (1996) investigated whether more accurate discrimination, in the statistical sense, could be achieved if models were built for separate subgroups within the population, and also looked at the rejection rates for these subpopulations. One of their clearest results was that, all other things being equal, subpopulation scorecards tended to reject fewer applicants than full population scorecards. At the same time the study proved that subpopulation scorecards did not always improve predictive accuracy - the subgroups needed to be sufficiently different.

Such differences in acceptance rates depending on whether segmented/non-segmented models are used, can be explained by a phenomenon referred to as the Yule-Simpson's paradox (Yule (1903); Simpson (1951)) in statistics. The paradox arises when the relationship between the outcome and predictor variables changes depending on the value of a third variable. Examples of the Yule-Simpson's paradox abound in medicine (Hand (1979); Hanley and Theriault (2000), Julious and Mullee (1994), Baker and Kramer (2001)), when, for example, the effect of treatment may appear unsuccessful for the whole population of patients, but successful if tested for male and female segments of the same population separately.

To illustrate the Yule-Simpson's paradox when applied to consumer credit, we consider the following example, which is totally hypothetical. Table 2.3 gives the breakdown of accepted applications by sex, employment status and outcome. The numbers in parenthesis are probabilities of being 'good' given employment status.

Table 2.3 Number of accepted cases by applicant's sex and employment status.

		Applicant's Sex					
		Male		Female		Combined (Marginal)	
	Outcome	Good	Bad	Good	Bad	Good	Bad
Employment status	Full-time	130 (0.62)	80	30 (0.75)	10	160 (0.64)	90
	Part-time	40 (0.57)	30	130 (0.72)	50	170 (0.68)	80

Whereas both male and female applicants with full-time jobs are better credit risks, the part-time status appears to be associated with more creditworthy behaviour, if no distinction between sexes is made. This may be the case when the use of gender is prohibited by legislation. So removing gender from the model will give misleading results. Another conclusion that follows from this example is that although women are better credit risks compared to men, they are not 'rewarded' for this in the 'sexless' scorecard.

Yule-Simpson's paradox arises because the combined probabilities are averages weighted by the fraction of each gender. Since a greater proportion of those with part-time status is women, more weight is given to women in the marginal probability for part-time status. On the contrary, the full-time marginal probability reflects the male dominance in this employment category. This situation can be easily remedied either by segmentation or stratification, but if there are legal restrictions on the use of sex in a scorecard, this does not appear possible.

Therefore, the inability to distinguish between different subgroups makes the lenders apply one generalised scale to all applicants, which is detrimental to both lenders and borrowers. Furthermore, if protected groups constitute the minority in the overall population, they will be assessed on the white male scale simply because their characteristics will be given less weight in the statistical analysis. It may sound paradoxical, but in order to eliminate discrimination in the social sense, it is necessary to discriminate in the statistical sense.

The problem of the legal interpretation of discrimination consists in trying to eliminate both direct and indirect discrimination, whereas in application to credit scoring these dual targets appear to be mutually exclusive. If the use of certain characteristics (direct discrimination) is prohibited, the protected groups are exposed to indirect discrimination. Since they are disadvantaged in comparison to other groups in the sense of having values of characteristics associated with high default probability, this state of the world is reflected in the outcome of the credit scoring decision. If on the other hand, the law wishes to combat the unequal distribution of credit (indirect discrimination), then the use of the relevant information should be allowed.

However, the legal argument (Banton (1994)) holds that the anti-discriminatory legislation is a means of achieving the equal treatment of persons. Using the prohibited information reinforces the distinction, which the law seeks to eliminate. Nevertheless, in our understanding, the law should distinguish between subjective and statistical discrimination, and in application to the latter, equal treatment could be interpreted as the right to receive an equally accurate assessment. Obviously, this would require the use of all relevant information, but this is regarded 'politically unpalatable' (Johnson (1992)).

Another possible solution was suggested by Hand (1998): by building separate models for sub-populations and keeping the same proportions of rejected applicants, one can ensure that credit is extended to 'protected' groups on an equal basis. The problem with this approach, as pointed out by Martin and Hill (2000), consists in general deterioration of credit quality in the long run, if more individuals are accepted with a higher likelihood of default. The resultant increase in costs and potential decrease in credit availability will have to be borne by all, including the protected groups.

In general, we believe that there is a need to re-assess the interpretation of 'equal treatment' and its relation to 'responsible lending'. Retention of the information, whilst unpalatable, may allow action to be taken to ensure that fairness is achieved. This makes the decision a policy action, rather than an artefact of the modelling. But until this action is taken, the existing anti-discriminatory legislation will continue to work against lenders and borrowers.

2.4 Data protection and credit referencing

2.4.1 The right on information privacy

The first part of the chapter explored the protection of human rights and the application of equality principles in the area of consumer credit. It has been shown that the level of anti-discrimination requirements varies from country to country, with the most stringent rules being imposed on U.S. lenders, while at the European Community level human rights have not been addressed until recently. But in contrast to other fundamental rights, the right to privacy has received very special attention both in the European Union and in the United States.

The distinction should be made between the concepts of privacy and data protection. Bennet (1992) defines privacy as 'the exclusiveness of the physical space around the individual, the autonomy of decision-making without outside interference, and the right to control the circulation of personal information'.

According to this definition data protection constitutes only one component of a broader term 'privacy' and can be regarded as equivalent to the concept of 'information privacy', the definition of which is given by Westin (1967): 'the claim of individuals, groups or institutions to determine for themselves when, how and to what extent information about them is communicated to others'.

The need to protect information privacy was first recognised in the late 1960s and was associated with the emergence of the 'post-industrial' or 'information society'. According to Bennet (1992), one of the main characteristics of the post-industrial society is the increasing value of information. Hence, the question of ownership of information becomes crucial, because it becomes an important resource, perhaps, more valuable than other resources, such as labour and capital.

Technological advances in data handling and communication bring more complicated relationship between the individual and those that have information about the individual. It is recognised that this information can give a certain power to a 'data holder' and this power can be abused. So from the public and legal perspective, it is necessary to re-assess relations between 'data subjects' and 'data holders' and to impose certain safeguards and controls.

2.4.2 Data protection in the EU

2.4.2.1 The national provisions before the Data Protection Directive

The level of importance ascribed to data protection can be demonstrated by the fact that even before the introduction of the European Directive on data protection in 1995, the majority of EU countries already had some data protection acts in place as shown in Table 2.4.

Table 2.4 The status of data protection legislature in the EU and the USA (before the Directive 95/46)

Country	**Legislation or action**	**Date of passage**
United States	Fair Credit Reporting Act	1970
Sweden	Data Act	1974
Germany	Data Protection Act	1977
France	Law on Informatics & Liberties	1978
Denmark	Private Registers Act	1978
Austria	Data Protection Act	1978
Luxembourg	Data Protection Act	1979
United Kingdom	Data Protection Act	1984
Finland	Personal Data File Act	1987
Ireland	Data Protection Act	1988
The Netherlands	Data Protection Act	1988
Belgium	Law on Privacy Protection in relation to the processing of Personal Data	1992
Portugal	Constitutional provisions	1976
Italy	No legislature in force	
Greece	No legislature in force	
Spain	Organic Law 5/1992	1992

Source: Bennett (1992) and European Commission, Directorate-General XV Internal market and financial services, http://europa.eu.int/comm/internal_market/en/media/dataprot/studies/index.htm

Bennet (1992) discovered a remarkable level of similarity in the general approach that MSs adopted in relation to data protection. Most MSs had statements of fair information policy based on the same general principles:

- Openness or transparency, which means that the very existence of record-keeping systems should be publicly known;
- The possibility for individual access and correction that generally allows the data subjects to verify any information that is kept on them;
- The principle of collection limitation that places some boundaries on the data held in databases following from the purpose of data collection;

- The principle of use limitation, based on the notion of relevance: data can only be used for the purpose it was collected for;
- The principle of disclosure, data should not be communicated externally (to third parties) without the consent of the data subject or legal authority;
- The security safeguards principle, which implies that appropriate measures should be taken to ensure confidentiality and to prevent destruction or modification of data.

But in spite of the similarity in general approach, the scope, detail and status of national laws differed considerably (Korff (1998)). In some countries the law covered manual data as well as automated data (France), whereas in some countries the law only covered the latter (the UK). Some made fundamental distinctions between the rules for public and private sectors (Germany), while others did not make such distinctions. As for the status of data protection, it ranged from the countries with specific provisions enshrined in their constitutions (the Netherlands, Portugal) to countries with no data protection legislation at all (Italy, Greece).

It was because of these divergences that the need for a directive arose, since it was recognised that such divergences might impede the smooth operation of the internal market.

2.4.2.2 The Data Protection Directive overview

Directive 95/46/EC of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (hereafter 'the Directive', European Parliament (1995)) was adopted on 24 October 1995 and required implementation not later than three years after this date.

The main objective of the Directive was to provide a working balance between the needs of data subjects and those of data holders by facilitating and encouraging the free movement of personal data while at the same time strictly protecting the privacy of the individual. This balancing approach can be traced throughout the Directive.

The Directive, similar to the other harmonisation measures, was based on the principle of the minimum harmonisation which implies that the MSs are not allowed to go below the specified threshold of privacy but they are given certain freedom in going above it. However, in the case of data protection the threshold was set quite high, which was inspired first, by the importance attached to information privacy and data protection. And second, by the technological developments that enabled the free flow of information across the borders but at the same time increased the danger of the breach of privacy.

The Directive's scope of application is very wide; it applies without any distinctions between the private and public sectors, between the format of data or the technology on which it is stored or transmitted, and between automated and manually structured data. Data processing is also defined very widely so that all data are caught from collection to destruction. The Directive imposes the criterion of legitimacy for processing, which follows from the balancing approach between the interests of 'data subjects' and the interests of 'data users'. The processing of personal data is considered to be legitimate either with the consent of the data subject, or resulting from the necessity of some important public interest, or if a balancing of the interests of data users and data subjects has shown that the interest of the former should prevail.

The requirement of legitimate processing gives the data subject the right to object to processing in some cases, including automated decision-making. However, it does not apply to situations where it is the data subject who wishes to enter into a contract. As might be expected, data subjects have rights of access to the data and a right to know the reasons, when an automated decision has been made which is unfavourable to them.

The Directive prohibits processing of sensitive information relating to such issues as racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership and the processing of data concerning health or sex life.

The Directive draws a distinction between the circumstances where the data are and are not collected directly from the data subject. In both cases the data subject should be informed of the purpose of the processing and given guarantees of fair processing. When the disclosure to third parties is envisaged, the data subject should be informed about it and the data subject has the right to refuse this use.

Member States must establish public authorities to be the supervisory bodies for the administration of the Directive within the territory of each MS, but in addition, a working party is established which includes a representative of the Commission and representatives of the MSs.

Finally, the Directive prohibits data transfers to those countries outside the EU which have an inadequate level of protection.

2.4.2.3 The implications for credit scoring.

On the surface, the Directive affected only the provisions of the credit contract (e.g. reasons for denial), but not the credit scoring practices. The clause concerning the sensitive data echoes, to a certain extent, the provisions of the anti-discrimination law. Such information – '**revealing the racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, health or sex life**'[7] -is not normally collected by lenders. However, special attention should be given to the word 'revealing', which potentially gives grounds to include some additional characteristics under the 'umbrella' of sensitive data.

At the national level the implementation of the Directive affected the activities of CRA that came under scrutiny of the national data protection authorities in some counties, e.g. Belgium, Greece and the UK. In the UK the Data Protection Registrar (now the Information Commissioner) used the Directive as the legal basis for resolving the long-standing dispute on the legitimacy of third party data in credit reports. Among other things, the British CRAs collect so-called third party data which includes:

- people with the same name, or a very similar name, living at the same address;
- other family members living at the same address;

[7] The bold is added.

- people with the same name, or a very similar name, who have in the past lived at the same address or the borrower's last address;
- other people who have in the past lived at the same address or the borrower's last address as part of the borrower's family;
- names of other people who have been listed on the Electoral Roll at the same address.

In September 2000 it was reported that the issue was resolved by the Working Party that was specifically set up for this purpose by the Information Commissioner, providing the Information Commissioner 'with a solution to her concerns, that also enabled the industry to extend credit without undue risk to the consumer or the lender' (Data Protection (2000a)). The solution involved the following measures:

- A shared surname and address will no longer be taken as an indication of a financial connection, e.g. parents and children are no longer automatically assumed to be formally connected.
- Customer requesting a copy of their credit file, will only see their own credit data and not that of any financially connected 'third party'.
- Individuals will be able to opt out of the automatic use of their financial partner's data.
- Household data will be used for fraud detection, and possibly, as a means of assessing over-commitment within a financial unit. (Data Protection (2000b)).

In France it was *credit scoring techniques* that were revised by the national authority. The significant number of complaints in the banking sector made the Commission National de l'Informatique et des Libertes (CNIL) carry out control missions to check the conditions of using credit scoring. It has emerged that for equal financial status, the criterion of nationality could enable lenders to discriminate between French nationals and nationals of another European Union country, or French nationals and nationals of a third country. This led the CNIL to issue a new recommendation of a general nature concerning **credit scoring** techniques. Such techniques **should have no relation to the nationality of clients**, as this would constitute unacceptable discrimination. Apart from nationality, this prohibition also affected the variable 'number of years at the last address', since it was considered to be related to nationality (CNIL (1998)). But as was mentioned earlier (Section 2.3.2) the French credit industry managed to win back the right to use 'nationality' and 'years at address', because it was established there was no intention to disadvantage other nationalities (Le Conseil d'Etat (2001)).

The Directive is very important, since it was the first Community measure related to the credit risk assessment practices. One important implication that follows from the Directive is that the lenders that are, in principle, subject to home country control, have to comply with data protection regulations of the country where the credit reference agency is established, if the lenders have to request the information on their 'foreign' applicants.

Although designed as a really tight harmonisation measure, the Directive suffers the same drawbacks of the harmonising measures that were outlined in Section 2.2. In general, although the Directive led the national authorities to consider privacy issues relating to credit referencing, the area of credit referencing still remains largely not harmonised, as will be shown in Section 2.4.4.

2.4.3 US data protection legislature

Although the Directive related to the MS of the European Union, the USA was affected by its provisions, because the Directive covered data transfers between the EU and external countries.

The importance that is attached to credit in US society can be demonstrated once again by the fact that Fair Credit Reporting Act (1970) was introduced earlier than a more general Privacy Act (1974). The **Fair Credit Reporting Act** (FCRA (1970)) still remains the main legislative tool for regulating the activities of credit reference agencies. The Act requires credit reference agencies to supply correct and complete information to businesses to use when evaluating a borrower's application.

Under the FCRA consumers have the following rights:

1) to receive copies of their credit reports;
2) to dispute the information held in the credit report. Both the CRA and the supplier of information are legally obligated to reinvestigate the dispute. If the

dispute is not resolved to the customer's satisfaction, the customer has a right to add a summary explanation to the credit report.

3) to know the name of anyone who received the credit report in the last year for most purposes or in the last two years for employment purposes;
4) to know specific reasons for being declined credit and the name and address of the CRA that supplied the credit report, provided the denial was based on information given by the CRA;
5) to "opt out" of inclusion on direct marketing lists.

The information is kept for seven years (apart from some minor exceptions). Particularly detailed reports, known as investigative reports, may be released only with notice to the consumer. The FCRA also requires that measures be taken to limit the dissemination of reports. Under the 1996 amendments to the Fair Credit Reporting Act, businesses can share certain consumer information with their affiliates, but they must first give customers the choice of opting out of the sharing.

In 1999 **The Gramm-Leach-Bliley Act** (GLBA (1999)), also known as the Financial Services Modernisation Act was passed which complements the limits and procedures on information-sharing already in place in the Fair Credit Reporting Act. The new act extends the scope of legislation to any entity engaged in financial activities and covers personally identifiable financial information about consumers. The law requires that a lender should provide consumers with a notice of a privacy policy and a chance to opt out of information-sharing with third parties.

It is difficult to say to what extent the new requirements set in the Gramm-Leach-Bliley Act were inspired by the EU data protection legislation, but after the Data Protection Directive came into force, the level of data protection in the USA was considered inadequate from the EU point of view. This launched negotiations between the USA and EU that lasted for two years. The Commission argued that the US took a sectoral approach to data protection which had produced 'a patchwork of federal and state laws and self-regulatory programmes' (European Commission (2000)). It assessed the adequacy of the level of protection afforded by the U.S. Fair Credit Reporting Act, in line with Article 25 of Directive 95/46/EC and concluded that it was not adequate.

However, in July 2000 The European Commission adopted a Decision determining that an arrangement put in place by the US Department of Commerce known as the 'safe harbor' provides adequate protection for personal data transferred from the EU. Under the 'safe harbor', US companies can voluntarily adhere to a set of data protection principles recognised by the Commission as providing adequate protection and thus meet the requirements of the Directive as regards transfers of data out of the EU. In many cases, for example under a specific statute such as the Fair Credit Reporting Act, which covers a number of situations where financial loss might occur (eg refusal of a loan), EU citizens will also have the option of taking the US organisation to court in the US.

Although the sectoral approach has been considered as a drawback by EC officials, the specific focus of US legislation on the particular area of application provides for effective information-sharing between lenders throughout the whole US territory. It will shown below that this is not the case in the EU.

2.4.4 Credit referencing

2.4.4.1 National differences in credit referencing

CRAs are important sources of information about the previous performance of a borrower. In many countries lenders agree to exchange information about their customers, this is done through credit reference agencies that act as information brokers. We have seen that credit regulations in general and anti-discriminatory legislation vary significantly from country to country. The same applies to the types of credit reference agencies and the scope of information they hold.

Jappeli and Pagano (2000) surveyed credit reference agencies in 49 countries and found that CRAs vary from country to country. In countries such as the USA, UK, Ireland and Sweden, credit information is provided by firms acting on a purely commercial basis (see Table 2.5). In other countries, the systems are managed by professional bodies or private organisations which are generally non-profit. Examples include the "Bureau Krediet Registratie" in the Netherlands, the "Schufa" in Germany, and the Belgian "Mutuelle d'information de L'union Professionnelle du credit". There are also countries where credit referencing is done by the Central Banks.

Table 2.5. Differences in types of CRA and information they hold

Country	Type of CRA	Starting Date	Type of Information
USA	Private	1890	Black and white
UK	Private	1960	Black and white
Austria	Private/public	1860/1986	Black and white
Belgium	Private/public	1987/1985	Black and white
Denmark	Private	1971	Black
Finland	Private	1900	Black
France	Public	1989	Black
Germany	Private/public	1927/1934	Black and white
Ireland	Private	1963	Black and white
Italy	Public	1990/1964	Black and white
Netherlands	Private	1965	Black and white
Portugal	Private/public	1977	Black and white
Spain	Private/public	1994/1983	Black and white
Sweden	Private	1890	Black and white

Source: Jappelli and Pagano (2000), Jentzsch (2003).

In countries with public CRAs there is a legal requirement for lenders to supply the information on the performance of their customers to the CRA, while in countries with private CRAs the exchange of information is done on a voluntary and reciprocal basis.

As regards the kind of data collected, CRAs are even more varied. Both 'white' (positive) and 'black' (negative)[8] information is collected in the Netherlands, Germany, UK, USA, whereas France and Denmark have 'black' data only. The level of detail also differs significantly from country to country. CRAs have the most complete picture of the borrowers' performance in the USA, whereas in Europe the most detailed CRA information is the UK.

[8] 'Black' data is usually understood as an account when three consecutive payments were missed following the definition given in the minute of understanding between the British Bankers Association and the Data Protection Registrar; 'white' data – all other information, including those accounts where payments are up to date (Howells, 1995);

The US CRAs collect four basic types of information:

- Identification and employment information – individual's name, birth date, Social Security number, employer, spouse's name. The CRA also may provide information about applicant's employment history, home ownership, income, and previous address, if a creditor requests this type of information.
- Payment history - accounts with different creditors, showing how much credit has been extended and whether it has been repaid on time. Related events, such as referral of an overdue account to a collection agency, may also be noted.
- Inquiries - a record of all creditors who have asked for individual's credit history within the past year, and a record of those persons or businesses requesting his/her credit history for employment purposes for the past two years.
- Public record information - bankruptcies, foreclosures, or tax liens.

In the US, apart from standard 'credit reports', CRAs hold so-called 'investigative consumer reports'. These are detailed reports that involve interviews with consumer's neighbours or acquaintances about his or her lifestyle, character, and reputation. These reports can be obtained only with the consent of the consumer and may be used in connection with insurance and employment applications.

A UK credit report contains the following information:

- electoral roll information that is used to check the borrower's address;
- county court judgements (CCJ) and administration orders for borrowers who failed to pay their debts on time, normally kept for 6 years;
- bankruptcy information, also kept for 6 years;
- data contributed by the lenders on the performance of individual accounts which is shared on a reciprocal basis, including both 'black' and 'white' data;
- a record of the searches or requests for the borrower's file from lenders;
- the Council of Mortgage Lenders' Repossession Register available only to the members of the Council of Mortgage Lenders, also kept for 6 years;
- a report from CIFAS, a credit industry fraud avoidance system, showing a fraud or an attempted fraud, available only to CIFAS members;
- "gone away" marker made by GAIN (Gone Away Information Network) when the borrower who is in arrears has moved without giving a new address;
- 'third party information' that was described earlier.

The level of detail of CRA data is the subject of constant debate, resulting from the need to balance the interests of different parties: data subjects' desire to keep their privacy, and the need of society to promote the policy of 'responsible lending', i.e. that credit should be given only to those that can repay it. There is a growing concern about the problem of overindebtedness. Therefore, the balance largely depends on the public perception of the benefits that follow from the use of the credit reference data.

2.4.4.2 The importance of CRA information in lending

Japelli and Pagano (2000) suggest that information-sharing is important for three main reasons:

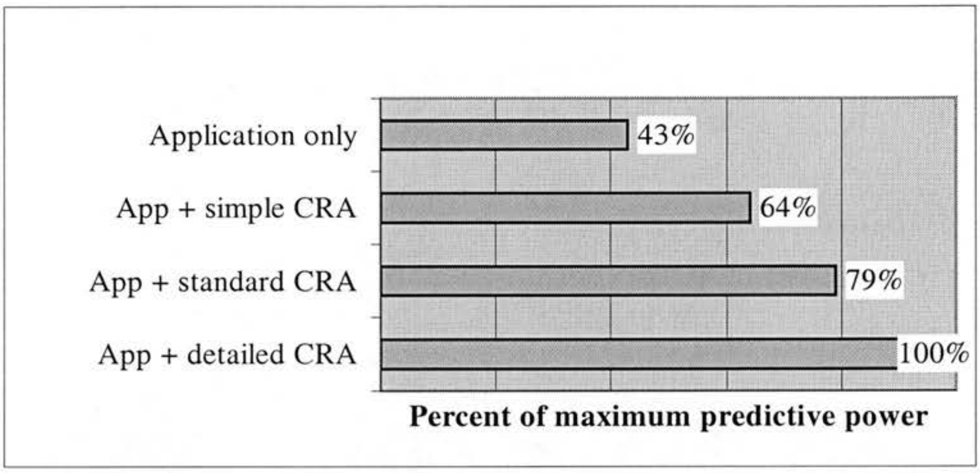
- CRAs improve the lenders' knowledge of their customers and therefore improve the prediction of their repayment performance;
- CRAs reduce the cost of obtaining the information that lenders could otherwise pass onto their customers;
- CRA information provision is a borrower disciplining device.

Following their analysis of information-sharing in 49 counties, Jappeli and Pagano concluded that information-sharing is positively correlated with the amount of lending and negatively correlated with the default rate.

There is more empirical evidence to support the importance of CRAs in lending in terms of the quality of prediction given by credit scoring models. Chandler and Johnson (1992) tested the contribution of data contained in US credit reports made to the predictive ability of scoring models. The relationship between the level of predictive power and the level of CRA data was studied under four scenarios with different levels of detail of CRA information. The predictive power of each model was measured by the Kolmogorov-Smirnov statistic that measures the ability of the scoring model to separate "good" and "bad" customers. Figure 2.1 presents the predictive power for different levels of CRA data that are compared on a percentage basis, with the highest Kolmogorov-Smirnov score scaled to 100 percent.

Their results suggested that the more detailed the CRA data, the better is prediction, the more likely it is that good customers will receive credit and that bad customers will be denied the use of credit. They concluded that placing limits on data retained and communicated by CRAs are 'anti-consumer'.

Figure 2.1 Predictive power of credit scoring models depending on the level of CRA report detail (Source: Chandler and Johnson (1992)).



Since the level of the detail and the amount of data held by CRAs varies from country to country, the significance of CRA data in credit scoring models will also be different. In countries with less detailed CRA information, the contribution of application information to predictive power would be expected to be greater than in countries with more detailed CRA information. Such asymmetries create additional difficulties in credit risk assessment on the European level.

Although, the Data Protection Directive has, to a certain extent, reduced the variations in data protection policies of the MS, credit referencing remains extremely diverse in terms of the legal status of the CRAs and the scope of information they hold. If the quality of prediction depends on the level of detail of CRA information, the applicants from counties with less detailed CRA will loose against others.

There is a number of practical questions on how the credit referencing should be carried out, when people apply for credit across national borders. Consider the situation when a Belgian national living in the Netherlands applies for a loan in

Germany. Should the German bank request the credit report from Belgium or the Netherlands or both? Where should the subsequent performance be reported to, if this customer is accepted? How will the level of detail of a credit report affect the chances of the Belgian national being accepted by the German bank? Will it be viewed as discrimination if his/her creditworthiness is assessed on a different basis as compared to the German national?

The proposal for a new EU Directive on consumer credit (European Commission (2002)) contains a section on 'central databases', and the MS will be obliged to set up and maintain 'negative' databases with the right to hold also 'positive' information on the financial performance of individuals. The Directive also intends to improve the mechanisms of credit report circulation between the MS. So hopefully, the questions above will be addressed shortly.

2.5 Conclusions

This Chapter has investigated the current state of integration of the European credit markets and provided the most detailed and comprehensive review of the restrictions on the use of information in credit scoring in the USA and EU. Such review has not appeared in previous studies.

It has been shown that in spite of the impressive progress towards the creation of an integrated market in financial services, the scale of divergence in national regulations affecting credit scoring remains significant. It means that even when the cross-border payments become cheap and efficient, the free flow of consumer credit across borders may be still obstructed, since the existing legal framework does not provide the basis for effective credit risk assessment across countries. But the EU authorities recognize the situation and work is in process to remedy it.

Still credit scoring techniques and, most important, the scope of information that can be used in credit scoring, are not specifically addressed either by existing or by proposed harmonisation measures. Thus in Europe information available for credit scoring becomes subject to general provisions of anti-discrimination and data protection law that cover a range of areas, not only consumer credit.

It has been demonstrated that Member States have significant freedom in interpretation and implementation of the EC law. Therefore, national anti-discrimination and data protection rules remain diverse. This may have the following implications for the lender's ability to assess the creditworthiness of a mixed population consisting of residents of several European countries.

First, it is not possible to obtain the same information for residents of different countries due to different anti-discrimination regulations and different levels of solvency data held by the national CRAs. Previous research has found that CRA information enhances the predictive power of credit scoring models and that the level of detail of CRA records is directly related to the quality of credit risk assessment. So residents of countries with less detailed CRA data will be subject to less accurate risk assessment.

Second, the principle of equal treatment can be interpreted in such a way that no distinction could be made between people of different nationalities. Since the law does not distinguish between subjective and objective discrimination, the prohibitive approach which is generally used in combating judgmental discrimination can be applied to credit scoring and therefore, the use of certain information may be regarded as illegal.

The analysis of previous studies has demonstrated that prohibition on information related to the probability of default impairs the predictive ability of credit scoring models and is detrimental for both lenders and borrowers. But until some clarification on equal treatment in credit scoring is given at the EU level, the danger remains that generic models may be regarded as the only legally acceptable ones.

The next two Chapters will investigate how much reduction in the predictive power of models can be expected if nationality cannot be used in credit risk modelling.

Chapter 3. Literature review on generic models in credit scoring

3.1 Introduction

The aim of credit scoring is to construct a classification rule that distinguishes between 'good' and 'bad' credit risks according to some pre-determined definition. This rule should classify the entire population of applicants for any particular credit product with maximum accuracy. However, the accuracy of classification may be affected when the overall population comprises heterogeneous subgroups. It is not uncommon in credit scoring to develop separate models for each subgroup.

With European integration in progress and developments in Internet banking, the scope of the target population is changing from the credit applicants of any given country to credit applicants of the EU or several European countries. The introduction of the Euro in 12 countries has removed the exchange rate risk faced by credit applicants, so one would expect more inter-country applicants for credit. According to Eurostat, the estimated population of the EU will be nearly 380 million people in 2005 (Table 3.1 gives the breakdown by country). Potentially this offers an extremely attractive possibility of market expansion for creditors, provided they can score not only the residents of their own country, but also applicants from the neighbouring Member States.

Several questions arise. First, how accurate will the classification be if the entire population of European applicants is scored with one model, and second, could the classification be improved by segmentation, i.e. by building individual models for different nations?

These questions follow from the legislative provisions underlying the EU harmonisation. Whilst segmentation can be valid from the statistical point of view, it may not necessarily be legal. As it was shown in Chapter 2, the use of 'nationality' in credit scoring models is open to interpretation by national authorities. If segmentation is not legal, it would be of interest to examine how much classification accuracy is lost if a generic rather than a national model is estimated.

Table 3.1. The EU population estimates in 2005
(Source: Eurostat (2003)).

Country	**Population projection (millions)**
Austria	8.12
Belgium	10.28
Denmark	5.42
Finland	5.24
France	60.32
Germany	82.99
Greece	10.66
Ireland	3.97
Italy	57.47
Luxembourg	0.46
The Netherlands	16.32
Portugal	10.14
Spain	39.65
Sweden	8.91
UK	60.25
EU-15	380.19

In a wider context this is a comparison of performance between generic models and customised ones. Thomas (2000) defines generic models as models developed on one or several populations or portfolios and applied to score a geographically or socio-economically different population/ populations. In contrast to this, customised models are developed on, and applied to, only one population.

This chapter presents a review of the previous research on generic models/ segmentation in credit scoring. Section 2 describes the existing approaches to modelling the changing environment, which is the case when creating one model for several countries. Section 3 investigates the conditions when linear models are insensitive to variations in coefficients, the so-called 'flat maximum effect'. Section 4 reviews the previous research on segmenting the populations in credit scoring. Sections 5 relates to the problem of multicollinearity which is a common feature in credit scoring datasets and one of the conditions for the flat maximum effect to hold. Section 6 concludes.

3.2 Concept drift and population drift

The question of selecting between generic and customised models depends on the level of the homogeneity / heterogeneity of the target population, which can evolve in time and space. We therefore start the discussion of generic models with some considerations of the underlying phenomenon – changes or differences in the population distribution – which may actually influence the performance of the generic model.

In recent years significant attention has been given to one aspect of the problem of varying credit risk patterns – population drift or changes in the characteristics of borrowers over time. In the machine-learning literature population drift is viewed as part of a more general phenomenon – concept drift, which may be caused by 'continuous change of the world and environment, or it may occur when the variables or the concepts depend on a certain (possibly unknown) *context*. The location can be seen as the context in which the data is collected – knowledge of the context (or a change of context) will aid the knowledge discovery process' (Taylor and Nakhaeizadeh (1997)). Concept drift can also refer to other changes, e.g. changes in the definitions of classes (good/bad) in supervised classification problems (Kelly et al. (1999)).

Two main problems arise when dealing with classification problems in a changing environment:

1) how to detect any difference or change;
2) how to react to identified differences. There are three ways of reacting: do nothing, develop a new scoring system, try to adjust the existing one.

The main methods of detecting and tracking changes are summarised by Taylor and Nakhaeizadeh (1997):

- Statistical Process Control (SPC) methods to detect changes and adapt or modify rules (Nakhaeizadeh et al. (1997)). The methodology assumes that the data are grouped into batches which are monitored over time;
- Machine-learning and statistical methods to detect contextual clues and react accordingly (Widmer (1997));

- Using a 'forgetting factor' whereby the rule is constantly updated using the most recently available observations (Nakhaeizadeh et al. (1996));
- Incremental (or on-line) learning with flexible 'forgetting operators' (Widmer and Kubat (1993), Widmer (1994))Widmer (1997).

The focus of the machine-learning approach is on the development of adaptive models with sequential estimation, where the classification rule changes each time a new observation or batches of observations are added to the data.

However, it is admitted that there is no general framework that would provide reliable guidelines for dealing with dynamic aspects and further research is required to develop some ideas presented in Nakhaeizadeh et al. (1997), Widmer (1997) and Mannila (1995) in order to produce simple and efficient solutions for robust modeling in a changing environment.

Countries can be considered as the varying context to react to. But in the area of credit scoring the investigation of concept drift is predominantly focused on the problems of population drift over time (Lucas (1992), Crook and Thomas (1992), Kelly et al. (1999)).

It is observed that the classification accuracy of credit scoring models deteriorates with time. The most common approach to the problem has been to adjust the cut-off and when this is no longer helpful, then re-develop the model.

Crook and Thomas (1992) investigated the effects of changing the cut-off. Two models for two different time periods were developed. It was found that the overall default rate, i.e. the prior probabilities of group membership differed between the time periods, and that default rates for each attribute of the predictor characteristics, were different as well. This resulted in changes in posterior probabilities of group membership, conditional on the attributes of the person, and was reflected in the outcome of classification – different proportions of applicants would have been accepted in different time periods. Even when the reject rate was held constant, the differences in proportions remained, suggesting that the decision for one and the same person might not be the same in different time periods. So the paper demonstrated that there were differences between classification rules developed in different time periods and that the cut-off changes do not guarantee the same decision for all applicants.

Lucas (1992) presented a method of monitoring population drift based on the difference between the actual and predicted good rate for all variables in the model. Confidence intervals were used to measuring the difference. The actual adjustment was done by manual modification of the predicted good rate for the poor performing attribute to match the actual good rate. Then the coefficients were re-estimated by

$$\boldsymbol{c} = (\boldsymbol{X})^{-1}\boldsymbol{g},$$

where c = coefficients vector,

X = conditional probabilities matrix: P(variable a | variable b),

g = good rates vector.

The formula can be applied to the variables that are not in the model to check if the model performance is consistent with their good rates. But this approach cannot be applied for comparison of models with different variable definitions, which may be the case when dealing with different countries.

Kelly et al. (1999) provide a formulation of a dynamic model which is capable of adjusting the parameters each time a new observation is added to a dataset. They distinguish the following types of changes in a population:

1. changes in class (good/bad) priors, $p(i)$, i.e. probability of belonging to class i;
2. changes in the class conditional distributions $p(x|i)$, i.e. the probability of having characteristic x given membership of class i,
3. changes in the posterior distributions of class membership $p(i|x)$, i.e. the probability of belonging to class i given the value of characteristic x (or in other words, the priors estimated by using the knowledge about x).

Kelly et al. (1999) argue that the classification accuracy is affected only by the third type of change, i.e. there should be no increase in error rate if the proportion of bads increases in the population, but the changes in an applicant's characteristics do not affect the classes differentially. However, they warn that this only holds for true distributions, so due to inaccuracies in estimation the classification performance may still be affected, if there are changes in class priors.

Furthermore, according to Kelly et al. (1999) the posterior probabilities remain unaltered when the class conditional distributions change only if x includes all variables influencing the class membership. But in real life this is seldom the case, and so all three types of changes may have an effect on classification accuracy.

The above studies approached the problem by attempting to trace the subtle differences that evolve over time. Contrary to this, generic scoring tries to capture the most salient features that are stable both in space and time. So it can be a possible solution for population drift in time as well as in space.

3.3 Flat maximum effect and generic models

The possibility of constructing a generic model for several different populations - or populations of different countries in our case – without significant loss in the quality of prediction follows from the 'curse of insensitivity' (Rapoport (1975), von Winterfeldt and Edwards (1982)) or flat maximum effect (Lovie and Lovie (1986)) which implies that when there is a large number of predictor variables, the predictive ability of linear models is insensitive to relatively large variations in regression weights. It follows that seemingly different linear models can give the same level of classification accuracy.

Lovie and Lovie (1986) specify the following conditions for the flat maximum effect to hold:

1. Predictor variables in the model should have the highest degree of association with the outcome variable – 'dominant' in Lovie and Lovie's notation.
2. The optimal subset of the predictor variables should be chosen to produce the best performance.
3. Predictor variables should be collinear.
4. The predictor variables should be coded in the same direction, preferably positive, in relation to the criterion variable (known outcome).
5. The criterion variable should be dichotomous to maximise the discrimination power of the model.

The empirical test of Lovie and Lovie's proposition was undertaken by Overstreet et al. (1992) who used ten credit-scoring models developed for different Southeastern US credit unions. A generic model was built by using the weighted average of coefficients from five models, and its predictive power was tested against the five remaining individual models.

It was found that the generic model performed relatively well, although the performance depended on the cut-off levels. E.g., for a cut-off of 300 (on a 1000-point scale) the generic model correctly identified 41.72% of bads in the validation sample whilst the customised models correctly predicted 58.22% of bads, whereas for a cut-off of 700, the corresponding percentages were 88.70% against 90.40%. Overall customised classification models were superior in absolute terms, but this was partially attributed to sample bias – the performance of the customised scorecards was tested on their development samples. So the researchers concluded the generic model performance was relatively good, and this together with benefits of lower development costs could make it a preferred option compared to customised models.

Further analysis was undertaken in Overstreet and Bradley (1996) on a more complete and updated dataset. Instead of averaging the coefficients, the development samples for customised models were pooled, and the model refitted. Due to a wider range of variables used in the estimation (CRA variables were included) the new generic model outperformed its 1992 counterpart. However, customised systems still demonstrated more accurate classification. In order to assess the relative benefits of the generic model, costs were assigned to misclassifications. The cost-benefit analysis favoured the customised models again.

Nevertheless, when the generic model was tested on the sample of loans not included into model training, its performance was comparable and in certain aspects (e.g. correct prediction of bankruptcies[8]) even superior to the in-house customised model which the Credit Union was using at that time. This was attributed to fact the Credit Union had expanded its customer base. Since these new segments of population had not been included into the training sample of the in-house model, it could not classify them correctly. However, the generic model performed well, in spite of the fact that it had not been developed on this particular population either. It was found that the predictive power of generic models was more robust over time compared to customised ones, thus probably, making the former more attractive in a long term perspective.

[8] Overstreet and Bradley (1996) distinguished between bankruptcies and charge-offs within 'bad' loans

3.4 Populations and subpopulations

There seems to be a general agreement that customised models outperform the generic ones in terms of classification accuracy (Overstreet et al. (1992), Overstreet and Bradley (1996); Chandler (1998); Makuch (1998); Platts and Howe (1997)) although this loss in accuracy may be compensated for by some other benefits, such as lower costs. On the other hand, Chandler (1998) argues that the predictive power of the CRA generic models is significant, and comparable to that of customised ones. However, these models are, in fact, systems of generic scorecards that often contain multiple models for different segments of customers.

Traditionally scorecards are built on the population of potential borrowers as a whole. However, there are situations when lenders would like to score different groups of customers within the population separately: e.g. new customers as opposed to the existing ones or young people (Wilkinson (1992)).

Banasik et al. (1996) investigated the feasibility of creating separate scorecards for different subpopulations, e.g. married / not married, have children / no children, retired/ not retired, 4 years or less at present address / 5 years or more, homeowner / tenant / other, etc.

The performance of the model built on the total sample was compared to models built on samples split into 2 or 3 segments and using different approaches to setting the cut-off levels. Creating separate scorecards is equivalent to estimating one equation with interactions of the segmented variable with all other variables in the model. E.g., for marital status the equation is

$$\boldsymbol{Y = b_0 + b_1 M + b_2 XM + b_3 XS},$$

where Y = predicted score,

M = {1 if married, 0 otherwise},

XM = {X if M=1; 0 otherwise},

XS = {X if M=0; 0 otherwise}.

This can be rewritten as two separate equations:

$\boldsymbol{Y - b_1 = b_0 + b_2 XM}$ for married

$\boldsymbol{Y = b_0 + b_3 XS}$ for not married.

It follows that each population segment will require a different cut-off level. The cut-offs were set in a following way:

- according to the rate of misclassification of goods. First, the error of classifying goods as bads was fixed at 10% for all scorecards. Second, a 10% error rate was applied to the whole population, but the subpopulation cut-offs were set so that the error rate was the same as in the full population scorecard for this population segment. Third, the sub-population cut-off matched the reject rate for this population segment in the full population scorecard when it misclassified 10% of the goods.
- according to the probability of default in the logistic regression,
- according to marginal good/bad ratios (3:1 and 5:1).

Overall, splits on 12 variables that gave 27 population segments were investigated and it was found that segmentation was not necessarily going to give the better classification accuracy. This was attributed to several reasons:

1) the necessity to have a separate cut-off for each segment with the danger of setting cut-off levels optimal for subpopulations which may not be optimal for the whole population;
2) a subpopulation represents a smaller sample compared to the whole population, which leads to higher variance, and therefore, less accurate estimation and prediction.

The overall conclusion of the paper is that the subpopulations need to be 'sufficiently different' to justify the development of separate scorecards. Wilkinson (1992) also warns against 'mixing dissimilar populations'. To gain benefits from separate scorecards, interaction of the segmented variable with other variables should be sufficiently significant, so that omitting interactions would result in a poorly-fitted model.

On the other hand, Makuch (1998) argues that a sample constructed in a comprehensive heterogeneous fashion (i.e. a sample should represent the population segments adequately) can produce a model that will apply to populations beyond those on which it was developed. This feature is referred to as 'cross-applicability'. An example is the FICO (Fair Isaac Credit Score) model which was developed originally to score credit cards, but then was successfully applied to mortgage scoring. The reason is that the posterior probabilities of class membership in the FICO development population hold for the mortgage applicants, which leads to 'cross-applicability'. However, if the unique relations captured by the model exist in the data, and those relationships are not shared by other populations then cross-applicability will not work.

The potential problems of generic cards were investigated by Staten (1999), Avery et al. (2000) and Barron et al. (2000). In the United States the majority of CRA data and delinquency survey data are reported for the whole country, sometimes (not always) with a breakdown for separate states, and very seldom for counties. Staten (1999) showed that there was a wide divergence in loan performance across states and counties and suggested that segmenting the data by states and counties would improve discrimination between good and bad loans.

The later paper by Avery et al. (2000) investigated this issue further and compared the predictive performance of models with county-level data to models with state-level data. The latter turned out to be less accurate than the former ones.

Barron et al. (2000) also argued for the incorporation of local economic data into credit-scoring models, as local factors showed significant correlations with credit scores, which is to be expected, since the applicants from the same location experience same local economic conditions. Lenders, while using CRA scores that are available from nation-wide CRAs, do not make adjustments for local conditions. It was shown that this may lead to unanticipated levels of risk in certain parts of the country.

Another set of problems which has been investigated by Barron et al (2000) is completeness of information and the representativeness of the sample that is used for developing the CRA scores. The analysis was conducted on a nation-wide sample of 3.4 million individuals that was stratified according to ZIP-code, and the variation on the ZIP-code level was then investigated.

Although this analysis relates to the US CRA generic models and the importance of incorporating regional/local information into CRA scores, the results of the analysis have implications for application models as well. The authors emphasised the fact that generic models may not necessarily perform equally well for all subgroups in the population.

This can result from:

1) Omitted variables bias: some of the variables not included into the analysis may have differential effects for different subgroups in the population;
2) Underrepresentation of certain subgroups in the sample which is used to develop a scorecard.

The authors concluded that CRA scores are affected by regional information and if this information is omitted this can lead to the distortion of results across different regions. They suggested two potential responses to the problem:

- developing subpopulation scorecards;
- adjusting scores according to the economic and regional context.

They also found that an adequate representation of different population segments in the development sample has an impact on the quality of prediction for these segments in the population. But the scale of the impact depends on how different their repayment behaviour is, compared to the population at large.

The only attempt to built a generic scorecard in the European context (Platts and Howe (1997)) also demonstrated the superiority of customised models. The analysis was conducted on retail credit databases for 5 countries that represented different European regions: UK, Germany (Northern Europe), Greece (emerging markets), Belgium (France/Belgium) and Italy (Southern Europe).

Three levels of models were built: one European scorecard, 5 country scorecards and 5 regional portfolio scorecards. The European model was built on 'around 20' global characteristics that were common to all 5 countries. For each attribute of a characteristic weights of evidence were calculated, that served as a basis for grouping attributes within a characteristic.

The country scorecards were developed on the same common characteristics, but the grouping and weighting of attributes was done separately for each country, so the resulting variables in the model and their estimated coefficients differed from country to country and from the European scorecard. Finally the portfolio scorecards that incorporated all available information for a particular portfolio were also included into the analysis.

The performance was measured by comparing the percentages of predicted bad debt achieved by each type of model, keeping the same acceptance rate. Both regional portfolio and country models proved to be superior to the Eurocard. The comparison to regional portfolio cards favoured the latter, since they incorporated more information than was available for the Eurocard development, but the country scorecards were developed on the same characteristics by re-classifying and re-weighting the Eurocard variables.

For all countries there was a marked difference in the bad debt improvement made by the country scorecard compared to the Eurocard. In the UK the difference was most pronounced: re-classifying and re-weighting gave a 12.5% improvement. Even in Belgium, where the country model was closest to the Euro model an improvement of 3.38% compared to the Eurocard was achieved.

The results obtained from the comparison of the Eurocard to country scorecards were explained by the authors as resulting from the following factors:

1) there were significant differences in applicant profiles across countries;
2) the predictive patterns in the data were significantly different in each country.

It was noted that differences that were present in the country models were averaged in the generic model, and that lead to inferior predictive performance.

The problem of mixing heterogeneous populations has been already referred to in Section 2.3.4 when the Yule-Simpson paradox was discussed. In the case of generic models the paradox will arise if there is a strong relationship between a country indicator and other predictor variables in the model. If there are only a few such variables the situation can be remedied by including the interaction terms between the country indicator and the affected variables. If there are too many such variables, then segmentation makes more sense. In any case the relations between variables in the model should be carefully examined to detect possible associations, including collinearity.

3.5 Multicollineartity

Multicollinearity can be generally defined as the situation when there are strong linear dependencies among the predictor variables. According to Belsley et al. (1980), multicollinearity exists if there is a high multiple correlation when one of the predictor variables is linearly regressed on the others.

The collinear variables do not provide much extra information above that already contained in the others. As a result it is difficult to infer the separate effect of such predictor variables on the response variable. The coefficients are unbiased, but they become less stable. Coefficient standard errors get large, so confidence intervals for parameter estimates are broad, reflecting the imprecision of estimation. Variables that seem to have weak separate effects, may have a strong combined effect. But this applies only to those variables in the model that are collinear.

In prediction problems, more interest is attached to the estimation of the combined effect of the variables rather than to the estimates of individual parameters. It was shown by Theil (1971) that specific linear combinations of estimated regression coefficients may be well determined even if this is not the case for individual coefficients. That is why multicollinearity is not often regarded as a problem in credit scoring, but there seems to be a general belief that one is better off without it. In the context of research, when interpretation of parameters is desirable, e.g. when comparing different countries, multicollinearity may present a problem.

Various methods of assessing multicollinearity have been proposed. One method is to examine pairwise measures of association – a correlation matrix for numerical data and χ^2 or Cramer's V for categorical data. However, this may not be sufficient, since it does not reveal dependencies between more than two variables.

A more useful approach is to examine variance inflation factors, VIF_i, which are the diagonal elements r_{ii} of the inverse correlation matrix R^{-1}, where i is predictor variable i. Their diagnostic value follows from the relation

$$VIF_i = \frac{1}{1 - R_i^2}$$

where R_i^2 is the multiple correlation coefficient of X_i regressed on the remaining predictor variables (Belsley et al. (1980)). Alternatively, one can look at tolerance which is the reciprocal of VIF.

Although, Belsley et al. (1980) argue that VIF cannot distinguish between several coexisting dependencies, and there is a lack of a meaningful boundary between what can be considered as high and low values, other authors (Allison (1999); Fox (1991)) believe that the performance of this diagnostic approach is quite satisfactory.

Belsley et al. (1980) suggest a condition index approach that is based on principal component analysis (PCA). PCA involves linear transformations that lead to a set of new uncorrelated variables. Although these new variables are not particularly useful for interpretation or prediction, they aid the understanding of the pattern of multicollinearity. The variances of components are eigenvalues. A condition index is the square root of the largest to the smallest eigenvalue. The value of 30 is often quoted as signalling serious multicolliearity (Freund and Littel (2000)).

If the presence of multicollinearity is established, a really serious problem is what to do about it. Multicollinearity is a property of specific characteristics of the data matrix X, and not of the statistical model. Since it is a data problem, and not a statistical one, the ideal solution (Belsley et al. (1980); Fox (1991)) would be to collect new data in such a way that a problem is avoided. Whilst this approach can be helpful in experimental studies, in credit scoring it is of no practical value.

Other approaches include model respecification, variable selection, biased estimation and using prior information about the model parameters (Fox (1991)).

Model respecification involves combining several predictor variables into an index. One of the ways of achieving this would be to include interaction terms for collinear variables. Or alternatively to develop several models for each level of the collinear variable, which is essentially the same as including interaction terms.

In variable selection the number of predictor variables in the model is reduced to a less correlated set by means of some procedure, most commonly forward or backward stepwise procedures, which add or delete variables one at a time on the basis of the specified significance level. Some authors (Fox (1991); Freund and Littel (2000)) believe that the stepwise selection methods may fail to return the optimal subset of predictor variables, so the best subsets technique is recommended as the more appropriate one. The latter technique examines all possible subsets of predictors and finds a specified number of models for all possible model sizes. However, in spite of the considerable advances in computer power, this procedure can be time-consuming and often not plausible with a large number of potential predictors. It also may suffer from overfitting by getting the best fit for the set of data used for training. Hosmer and Lemeshow (2000) argue that stepwise selection procedures are useful and effective data analysis tools capable of producing models equivalent to models with purposeful selection of variables.

Biased estimation consists of using techniques like ridge regression, where a small amount of bias in parameter estimates is traded for a significant reduction in parameter variance. However, ridge regression involves the selection of an arbitrary "ridge constant". When it is large, the bias is large and the variance is small. To select a reasonably good constant one needs to have some knowledge about the unknown parameters that are being estimated. So the procedure should be applied with great caution.

Biased estimation can be viewed as part of the next approach – introduction of prior information about the parameters. There are several other ways to do this, including the formal Bayesian analysis (Belsley et al. (1980); Theil (1971)). But the basic idea can be demonstrated by the following example derived from the illustration offered by Fox (1991).

Suppose we would like to estimate the parameters in the model

$$\log(\frac{\boldsymbol{p}}{1-\boldsymbol{p}}) = \beta_0 + \beta_1 \boldsymbol{x}_1 + \beta_2 \boldsymbol{x}_2$$

where p is the posterior probability of being good, x_1 is the applicant's age and x_2 is the spouse's age, the variables that are normally highly correlated. If we have reasons to believe that $\beta_1 = \beta_2$, we can denote the common quantity as β^* and fit the model

$$\log(\frac{\boldsymbol{p}}{1-\boldsymbol{p}}) = \beta_0 + \beta^*(\boldsymbol{x}_1 + \boldsymbol{x}_2)$$

The approaches presented here generally overlap, but none offers an ideal solution. Judgement and thought are required when dealing with multicollinearity.

3.6 Conclusions

This Chapter provided a review of the previous research on modelling changing environment and generic models. A number of studies in credit scoring have investigated the problem of population drift or changes in population over time. This problem has been addressed by adapting the models to capture subtle differences that evolve over time. Differences presented by credit applicants from several countries can be viewed as a manifestation of population drift across space. Generic scoring addresses this type of population drift by modelling the most salient features of a population which makes generic models more stable than customised ones, not only in space, but in time also.

Generic models utilise a property of linear models - flat maximum effect. It implies that relatively large variations in coefficients have little effect on the quality of prediction. Such robustness of linear models is especially evident when major predictors are included into analysis and with collinear data. Collinearity may present a problem when interpretation of parameter estimates is required.

The majority of previous studies on generic models show that although generic models have an advantage of lower costs of developing and maintaining, the predictive accuracy of generic models is inferior to that of customised/segmented ones. But it has been also shown that segmentation will only enhance the predictive ability of models when the segments in overall population are significantly different and relatively large in size. Otherwise the segmented model may demonstrate increased variability.

An appropriate representation of different population segments in a sample used for training a generic model is an important requirement ensuring that the model works well for all subpopulations. To be applicable to a wider range of different portfolios, the development population has to be diverse.

The next Chapter will present a comparison of the performance of customised models built for three European countries (Belgium, Germany, the Netherlands) to that of the generic model.

Chapter 4. Generic scoring using logistic regression

4.1 Introduction

The objective of this Chapter is to test the proposition that the 'flat maximum effect' can compensate for the differences in heterogeneous datasets, and a generic model can produce the classification performance comparable to the customised models.

A logistic regression model was fitted to three datasets from different countries (Belgium, Germany, the Netherlands) separately. Then the generic model was built on the same three datasets aggregated in one, and its predictive performance was compared to that of national models. It was found that the predictive accuracy of the generic model is close to that of national ones. This is attributed to the 'flat maximum effect' and the socio-economic proximity of the countries used in the analysis. The four models described above were developed on a set of variables that was common to all three countries.

In addition, three models were developed on a full set of characteristics that was available for each country. The resulting improvement in predictive performance compared to the generic model depended on the scope of additional information but was significant for all three counties.

This suggests that generic scoring is possible, at least for some regions of Europe, and that the current limitations of its application stem from discrepancies in data collection rather than from the loss of predictive accuracy due to re-weighting of variables in national models.

The analysis described in this Chapter is limited to application characteristics. The decision to look at application characteristics only can be justified by the following considerations. The application characteristics are mainly demographic ones, and the variables themselves and their measurement scales are similar across the countries. So it was possible to select a reasonable set of variables that was common to all three countries.

It may be, though, that the commonality disguises the difference contained within these measures. Differences in applicants' characteristics may be due to the concepts they are meant to measure. This variation is influenced by the complex blend of social, economic and cultural factors, and is beyond the creditor's control.

The CRA variables, on the contrary, appear to be very different for each country. Their distinctiveness stems, however, from the differences in measurement scales rather than in the concepts themselves. This means they can be harmonised, i.e. they can be measured with one common scale, like in the USA, where the CRA characteristics are harmonised, and generic models can cover the whole country (Chandler (1998)).

The attempt, though, to achieve harmonisation would take considerable time and effort. The European Commission has been considering it for a number of years, and so far has not achieved consensus. It would be inappropriate within the context of the current research to try to guess the outcome of the EC's deliberations. Any selection may subsequently be regarded as arbitrary in the light of the final result of the EC's harmonisation.

For similar reasons (Section 4.2 will outline some problems with harmonised statistical data at the European level) the inclusion of economic variables was not pursued in this analysis. Besides, the length of time (25 months) represented by the data in the analysis is too short to investigate the impact of the changing economy on the customer's behaviour.

The next section provides some background information about the countries used in the analysis. Section 3 presents the development and performance of the national models, including data description, definitions of good/bad, model specification and prediction results. Section 4 covers the aspects of development of the generic model and compares its performance to the national ones. Differences between applications accepted by different models are also investigated. Then the importance of additional information is examined. Finally, Section 5 concludes.

4.2 Background information on countries used in the analysis

The countries used in the analysis (Belgium, the Netherlands, Germany) are neighbours, they are the EC founding members and have joined the Euro from the moment it was introduced in 2002. Tables 4.1 – 4.3 present some basic demographic and economic indicators that show some similarities and differences between the countries.

Table 4.1 Population and demographics. Belgium, the Netherlands, Germany (2000). Source: Eurostat (2003)

	Pop-n as % of EU-15	**Pop-n increase, 1991=100**	**Marriages per 1000 people**	**Divorces per 1000 people**	**% aged 65 and over**
EU-15	100	103	5.1	1.9	16
Min	0.1 (Lux)	101.6 (Italy)	4.3 (Greece)	0.7 (Italy)	11 (Ireland)
Max	21.8 (Germany)	113.3 (Lux)	7.2 (Denmark)	2.7 (Fin, Denmark)	18 (Italy)
Belgium	2.7	102.5	4.4	2.6	17
The Netherlands	4.2	105.7	5.5	2.2	14
Germany	21.8	103	5.1	2.4	16

In terms of population Germany is the largest EU country, whereas Belgium is the second smallest. The Netherlands is the sixth largest country in EU, and with a land area only slightly bigger than Belgium, the Netherlands is the most densely populated EU country. The populations of Belgium and Germany increased at the average EU rate, the increase in the Netherlands was more rapid. This is in line with the proportion of older people: in the Netherlands it is lower than in the other two countries. The marriage rate is the lowest in Belgium, and highest in the Netherlands, with Germany being exactly at the EU-15 average. As for divorces, all three country are above the European mean, with lowest number being in the Netherlands and highest in Belgium. It looks like Belgium has a problematic situation with demographics, being quite small, it has in addition a low rate of population increase and low number of marriages, coupled with a high number of divorces and a large proportion of older people.

In terms of Gross Domestic Product (Table 4.2) all three countries have quite high GDP per capita, above the EU average, with the highest being in Netherlands and the lowest in Germany. However, the GDP growth rate is below the European mean, but the Netherlands shows higher growth compared to the other two countries. This indicates that all three countries have well-developed mature economies that do not grow at a very fast rate. The Netherlands also has the lowest unemployment rate in Europe, but one of the highest inflation rates. The Dutch top the EU list of the percentage of part-time jobs, it is twice as much as in Belgium or Germany. Overall, it is possible to conclude that the Netherlands stands somewhat apart from Germany and Belgium in terms of economic indicators.

Table 4.2 Economy. Belgium, the Netherlands, Germany (2001).

Source: Eurostat (2003)

	GDP per capita, EU-15= 100	**GDP growth rate, % change on previous year**	**Unemployment rate, %**	**Inflation rate, %**	**Part-time jobs, % of all jobs**
EU-15	100	1.5	7.3	2.3	17.9
Min	82.4 (Greece)	0.6 (Germany)	2.4 (NL)	1.2 (UK)	4 (Greece)
Max	197.4 (Lux)	5.7 (Ireland)	10.6 (Spain)	5.1 (NL)	42.2 (NL)
Belgium	106.5	0.8	6.6	2.4	18.4
The Netherlands	112.3	1.3	2.4	5.1	42.2
Germany	104.1	0.6	7.7	2.4	20.3

In terms of strength of credit institutions all three countries show relatively high indicators (Table 4.3). Credit institutions in Germany have the largest balance sheet totals in the EU, as well the as the highest volume of interest receivable. Belgium and the Netherlands show more modest results.

With the same numbers recalculated per inhabitant, the EU average is not particularly informative, since Luxembourg clearly represents an outlier. For balance sheet totals per inhabitant the Netherlands show a median result, with Belgium and Germany being above it. For the volume of interest receivable per inhabitant, all three countries are above the median, with the Netherlands showing the lowest number (out of three countries) and Belgium – the highest.

Table 4.3 Credit institutions. Belgium, the Netherlands, Germany (2001). Source: Eurostat (2003)

	Balance sheet total, Mln EUR	**Interest receivable and similar income, Mln EUR**
EU-15 (average)	1,562,172	84,685
Min	139,939 (Finland)	6,466 (Finland)
Max	7,037,504 (Germany)	358,962 (Germany)
Belgium	738,123	69,908
The Netherlands	1,018,788	54,689
Germany	7,037,504	358,962
	Balance sheet total per inhabitant, EUR	**Interest receivable and similar income per inhabitant, EUR**
EU-15 (average)	929,774	50,403
Min	13,942 (Greece)	1,151 (Greece)
Median	63,726 (Netherlands)	3,392 (Austria)
Max	1,475,510 (Luxembourg)	113,651 (Luxembourg)
Belgium	71,921	6,812
The Netherlands	63,726	3,421
Germany	85,552	4,364

Unfortunately, harmonised statistical information on consumer credit in Europe is limited and patchy. This reflects the level of integration in retail financial services, demonstrating that consumer lending markets are still fragmented. Eurostat does not publish data on consumer credit. The OECD provides some data on the household debt but not for all European countries. The IMF does not have a separate consumer credit category. The Centre for European Policy Studies and the European Parliamentary Financial Services Forum publications contain some statistics, but mainly the EU aggregates with occasional country break-down, and again not for all EU countries. The European Credit Research Institute (ECRI) has published a report on consumer credit statistics (Krastanova (2003)), but not all statistics cover all EU countries. Problems with data on consumer credit were noted by ECRI (Guardia (2000)) and by the European Commission itself (European Commission (1995)).

Le Codran de COFIDIS, a survey of the consumer credit market in Europe, published by the European telephone credit company (COFIDIS (2003)) is probably the most detailed free source of comparable cross-country information, but only for eight countries that include Belgium and Germany, but unfortunately not the Netherlands.

According to Le Codran de COFIDIS (COFIDIS (2003)) in 2002 the total consumer credit outstanding for the eight countries (France, UK, Germany, Italy, Spain, Belgium, Portugal and Greece) amounted to 725.9 billion EUR against 688.8 billion a year earlier. This represents a 5.4% increase, which is one point less than that recorded in 2001.

COFIDIS distinguishes between three different types in the European credit market:

- The markets of Northern Europe (France, Germany and Belgium), highly developed and mature. With consumer credit having expanded significantly over the last ten years, these markets now develop at a more modest rate and do not offer the same growth prospects.
- The markets of Southern Europe (Italy, Spain and Greece), which have developed more slowly, but now appear to be more dynamic. They currently offer greater growth potential than the markets of Northern Europe.
- The UK market, that is classified as 'atypical'. Already highly developed and unparalleled in Europe, it continues to post growth rates similar to those seen in Southern Europe.

Table 4.4 Outstanding consumer credit.

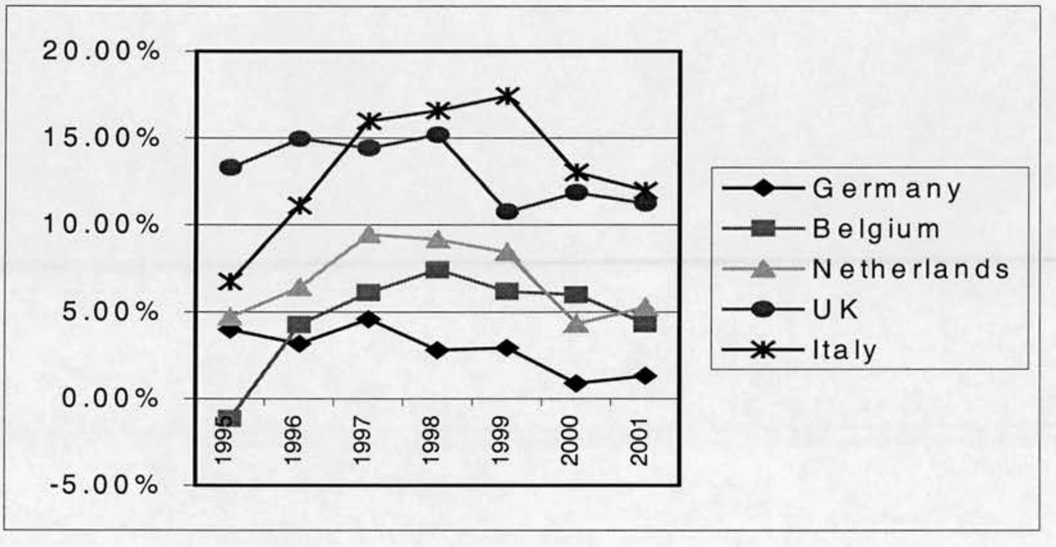
Source: COFIDIS and De Nederlandsche Bank.

Total outstanding consumer credit, bln EUR									
	Germany	**Belgium**	**Nether-lands**	**France**	**UK**	**Italy**	**Greece**	**Portu gal**	**Spain**
1993	176.9	10	8.6	58.7	83.5	16.8	0.5	-	30.8
1994	186.1	9	9.0	60.2	90.9	17.4	0.7	-	32
1995	189.5	9.3	9.4	63.8	106.7	18.3	1.4	-	36.7
1996	198.8	9.6	9.9	69.3	121.2	20.4	1.9	-	38.6
1997	204.3	10.2	10.9	74.9	138	23.8	2.1	5	42.4
1998	216.7	11.1	12.3	80.8	159.8	28.5	3	6.2	52.3
1999	215.7	11.9	12.8	89.5	185.2	32.3	3.9	6.8	54.6
2000	222.6	12.2	13.8	97.8	186.2	38.5	5.5	8.3	58.6
2001	222.4	13.2	13.9	103	221.8	41	7.9	8	62.4
2002	224.3	13.5	15.0	105.7	252.9	45.2	9.8	7.9	66.7
Total outstanding consumer credit per inhabitant, EUR									
	Germany	**Belgium**	**Nether-lands**	**France**	**UK**	**Italy**	**Greece**	**Portu gal**	**Spain**
1993	2,175	990	562	1,020	1,432	294	48	-	785
1994	2,282	888	585	1,042	1,554	304	66	-	814
1995	2,316	917	604	1,101	1,818	319	131	-	932
1996	2,424	944	636	1,192	2,058	355	177	-	978
1997	2,490	1,001	694	1,285	2,335	413	194	495	1,071
1998	2,641	1,087	778	1,381	2,691	495	276	611	1,317
1999	2,625	1,162	810	1,523	3,106	560	358	667	1,366
2000	2,706	1,189	865	1,656	3,110	666	503	809	1,451
2001	2,698	1,280	863	1,736	3,750	719	719	774	1,528
2002	2,718	1,304	924	1,773	4,263	789	889	759	1,605

The market is dominated by the countries of Northern Europe. The three large countries of Northern Europe (France, UK and Germany) represent more than 80% of total outstanding debt. The UK continues to be the number 1 market for consumer credit. If the UK figures are excluded, the growth of the credit market in 2002 drops to 3.3% (COFIDIS (2003)).

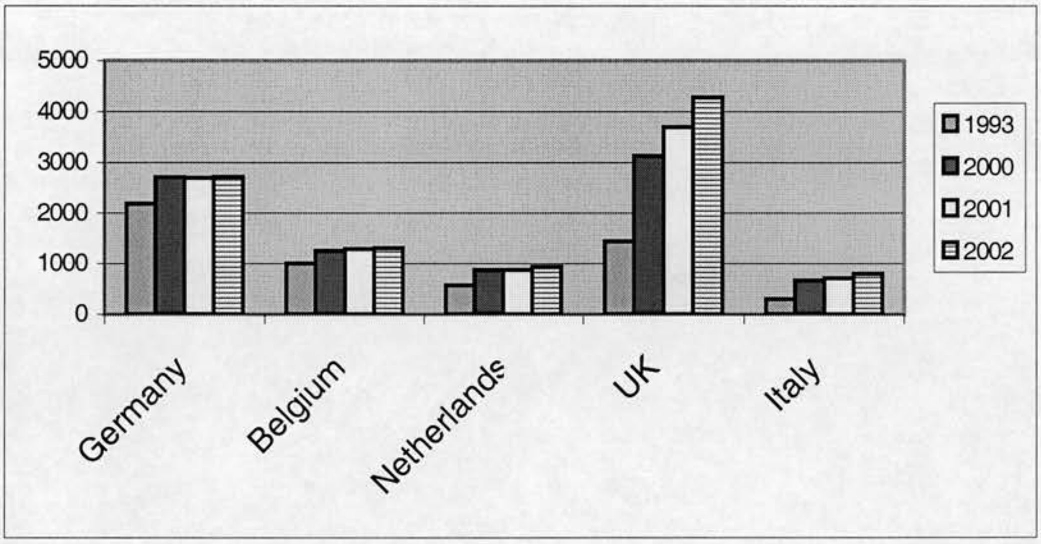
Table 4.4 and Figures 4.1-4.2 provide some statistics on consumer credit drawn from COFIDIS (2003) and De Nederlandsche Bank (2003). As for the total outstanding consumer credit (Table 4.4), the Netherlands follows quite closely the pattern presented by Belgium. Both countries are among the smallest in terms of the absolute credit volume. On the contrary, Germany is the second largest market in Europe after the UK, both in terms of the total volume and in terms of credit per inhabitant. As for credit per inhabitant, Belgium approaches the level of France, the Netherlands appears to be on the borderline between mature economies of Northern Europe and developing markets of Southern Europe. However, one should not forget that the Netherlands is the most densely populated country in Europe.

Figure 4.1 Annual growth rate of the total outstanding consumer credit. % increase on previous year (3-point moving average).



But the annual growth rate (Figure 4.1) for Belgium, Netherlands and Germany is notably lower than in the UK and Italy, which represents Southern Europe in this Figure, so that the comparison can be made between three market segments according to COFIDIS classification. In spite of dominating the market in terms of volume, Germany shows the lowest growth rate, but together with the Netherlands it indicates an upward trend in the end, although obviously it cannot be taken as an indication that this upward trend will continue into the future.

Figure 4.2 Average outstanding credit per inhabitant. EUR.



The average amount of credit per person (Figure 4.2) is the second largest for Germany, but the year-to-year change is modest, indicating the market maturity. Belgium and the Netherlands show a similar pattern, although the amount of credit is nearly half the value of Germany. Overall, the information presented above gives grounds to conclude that the Netherlands can be classified as a country of 'Northern Europe' according to COFIDIS market segmentation.

Figure 4.3 Consumer credit as a percentage of household disposable income (2001-2002). Source: ECRI

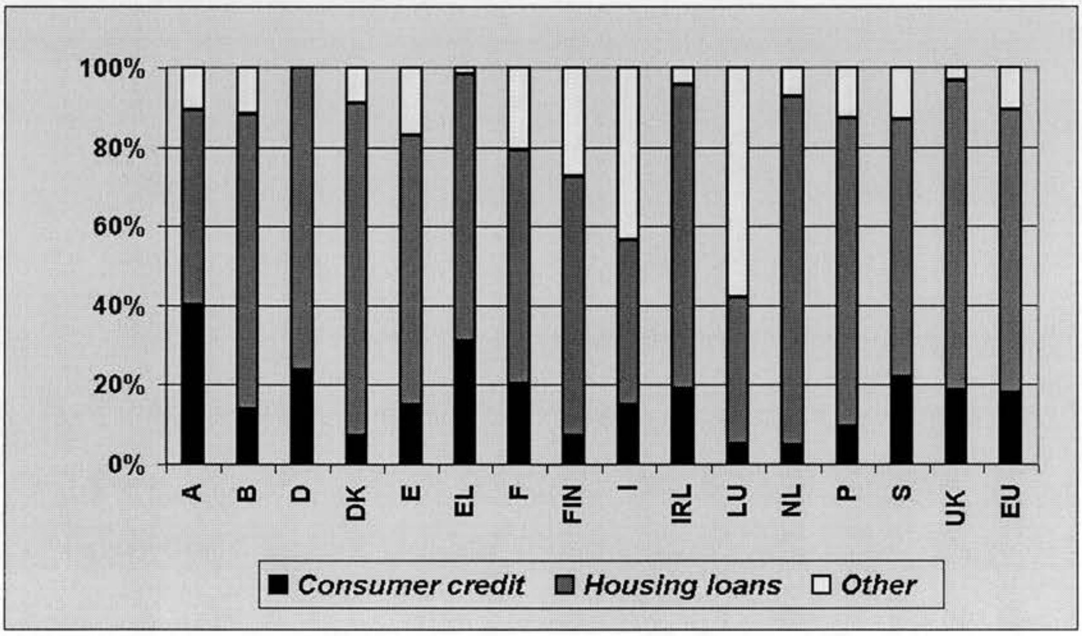


Figure 4.3 gives the breakdown by the type of loans as a percentage of the household's disposable income. It is notable that the Dutch spend more money on mortgages than probably any other European country. The Dutch Office of Statistics also reported a significant increase in mortgages in 2001-2003 due to the low interest rate, which was 4.5 percent in 2003, its lowest level since the mid-fifties (CBS (2003)). However, the Dutch spend a far smaller proportion of their money on consumer credit in comparison to Belgians and Germans.

In general, it is possible to conclude that the level of similarity/difference between the three countries analysed in this thesis depends on the perspective. If taken on their own they show some significant differences, if placed against the other European countries, then similarities become more evident. The Netherlands in a number of ways appears to differ from Belgium and Germany. But overall the three countries have enough in common to allow for a successful development of a generic model.

4.3. National models

4.3.1. Data description and definitions

The data for analysis was provided by a major international credit scoring consultancy and relate to the same retail card issue to applicants from 3 European countries: Belgium, Germany and the Netherlands. The card was managed by one bank and was offered through a range of participating stores to buy 'white' durable goods. A complete list of items that could be bought is given in Appendix, Table A10. Each account was given a credit limit, thus several purchases could be made. The populations of card-applicants differed in size and the period of time for which the performance was recorded. This is shown in Table 4.5.

Table 4.5 Populations of card applicants in three countries

	Belgium	The Netherlands	Germany
No of applications	108517	566960	894251
No of periods (months)	26 from 01/10/1998 to 30/11/2000	35 from 01/01/1998 to 30/11/2000	68 from 01/04/1995 to 30/11/2000
Characteristics	43	85	56

The acceptance period of 14 months was selected to cover the period from November 1998 until December 1999 for all three countries. The applications received during this period only were considered. Then their performance was observed until November 2000. The observation period, during which the performance was recorded, ranged from 12 months (for accounts that joined in December 1999) to 25 months (for applications received in November 1998). The applications received during the year of 2000 were not included in the analysis, since the credit type is revolving and the borrowers need to stay on the books at least for 10 months before any judgements about their creditworthiness could be made (Lewis (1992)). The applicants that were accepted by the creditor, but did not take up the offer were dropped as well. This group was small for all three countries.

The definition of bad was chosen to be 'at least 2 months in arrears' at any time during the observation period. The more traditional definition of bads of being 'at least 3 months in arrears' did not give sufficient numbers of bads to allow for effective classification. The remaining accounts were treated as goods[9]. The resulting samples were as shown in Table 4.6. The rejected applications were not used in the analysis for reasons outlined in Section 1.2.

Table 4.6 Samples used in the analysis.

	Belgium		**The Netherlands**		**Germany**	
	Frequency	*Percent*	*Frequency*	*Percent*	*Frequency*	*Percent*
Bad	3090	6.98%	11213	8.55%	8909	6.50%
Good	23200	52.42%	71024	54.13%	66939	48.81%
Rejected	17966	40.60%	48953	37.32%	61288	44.69%
Total	44256	100.00%	131190	100.00%	137136	100.00%
Odds of being good	7.5108		6.3341		7.5136	

[9] For a number of accounts it was difficult to decide whether they should be classified as good or bad, e.g. the accounts that remained 'Good' throughout the observation period, but showed 'Nil balance' or '1 month in arrears' in the last month. Such indeterminate accounts were removed from the samples. This is a standard practice in credit scoring to improve the separation between good and bad. (Lewis (1992); Thomas et al. (2002))

All three samples of accepted applications were split randomly into training (70%) and holdout (30%) datasets.

The list of characteristics collected in each country was different, they are given in Table A1 in Appendix. Those available are marked with 'X'. It was possible to select 16 characteristics that were collected for all three countries (Table 4.7) and they were used in modelling (given in bold in Table A1).

Table 4.7 Characteristics used in analysis

No	Characteristic	No	Characteristic
1	Home telephone	9	Employer's phone
2	Residential status	10	Card insurance
3	Marital status	11	Credit insurance
4	Occupation (Full-time, part-time, self-employed, etc.)	12	Number of dependants
5	Age	13	Spouse age
6	Time at address since 18 years old	14	Goods code
7	Time in employment	15	Goods price
8	Type of business (Manufacturing, banking, catering, etc.)	16	Payment date

4.3.2. Coarse-classification

For each attribute of categorical variables the weights of evidence (WOE) were calculated:

$w_{ij} = \log (g_{ij}B_j/b_{ij}G_j)$,

where g_{ij} (b_{ij}) are the corresponding numbers of goods and bads within the attribute i of characteristic j, G_j (B_j) are total numbers of good/bad in characteristic j in the sample.

The attributes with similar weights of evidence were grouped together into one variable, which is the standard practice in the industry and called coarse-classification. Judgements about similarity were done subjectively on the basis of visual inspections of plots of WOE, histograms and prior knowledge, where applicable. E.g., characteristic 'Goods code' originally was split into a lot of attributes, some of them contained very few observations. Such small categories were grouped with a larger category of a similar meaning, for example, 'Ladies boutique' and 'Ladies shoes' were grouped together with 'Ladies clothes'. In other cases grouping was possible only on the basis of WOE, since there was not enough explanation of the meaning available, e.g. different types of 'Benefit' of characteristic 'Business type' in the Netherlands.

The same approach was applied to continuous variables that were first divided into 5%-percentiles, which were then grouped together according to the weights of evidence.

Figures 4.4-4.5 show examples of the weights of evidence for two characteristics for each country, and the groupings that were chosen. Tables A3 to A16 in the Appendix present WOE by country for the remaining 14 characteristics that were used in the analysis.

The differences that were identified at that stage:

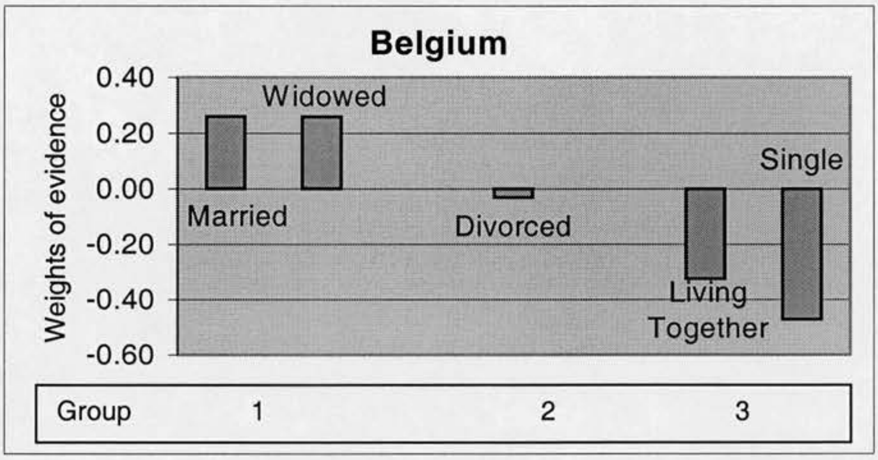
- For categorical characteristics in some countries there existed attributes that were not present in other countries (e.g. 'Living Together Registered' was a separate category of 'Marital Status' for the Netherlands but did not exist for Belgium or Germany; or 'Living on Boat' was available for the Netherlands but not for the other two in 'Residential Status').
- The distributions of good rate differed across the characteristics between the countries, which resulted in different coarse-classifications of attributes.

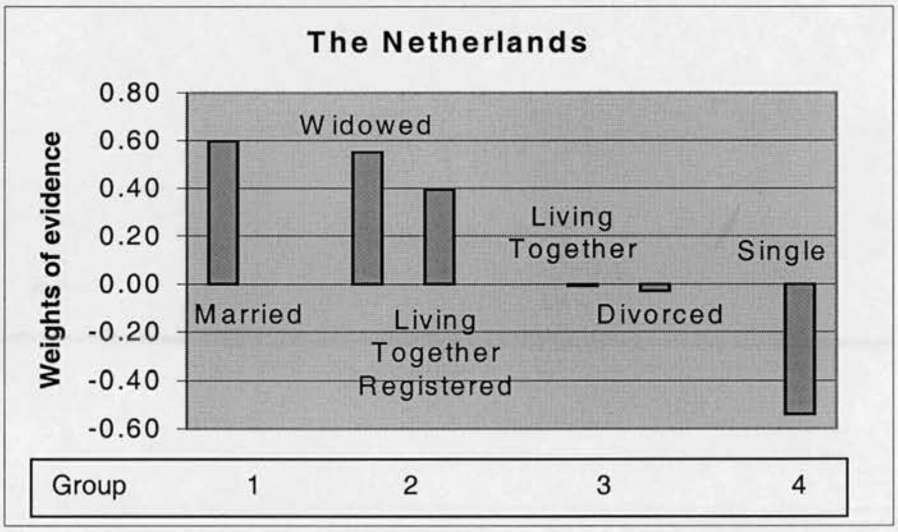
In other words, the differences were observed in

1) prior distributions of class membership $p(i)$, although this referred only to the Netherlands (from Table 4.6 - probability of being good is 0.88 for Belgium and Germany, and 0.80 for the Netherlands);
2) posterior distributions $p(i|x)$ (Figures 4.4-4.5).

Therefore according to the argument by Kelly et al. (1999), section 3.2, there was evidence of population 'drift' (albeit between countries rather than over time) and one might expect that differences in prior/posterior distributions will affect the classification accuracy of the models.

Figure 4.4. Coarse-classification of 'Marital Status' by country





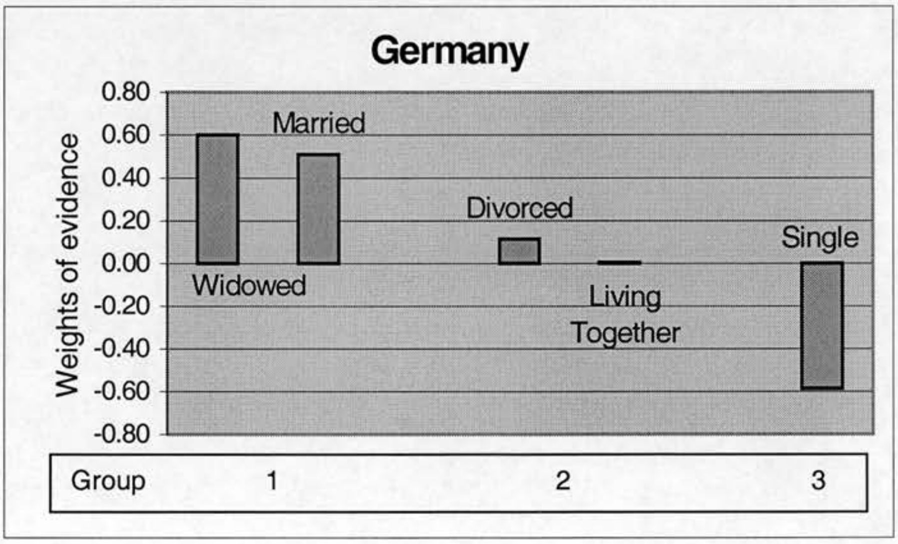
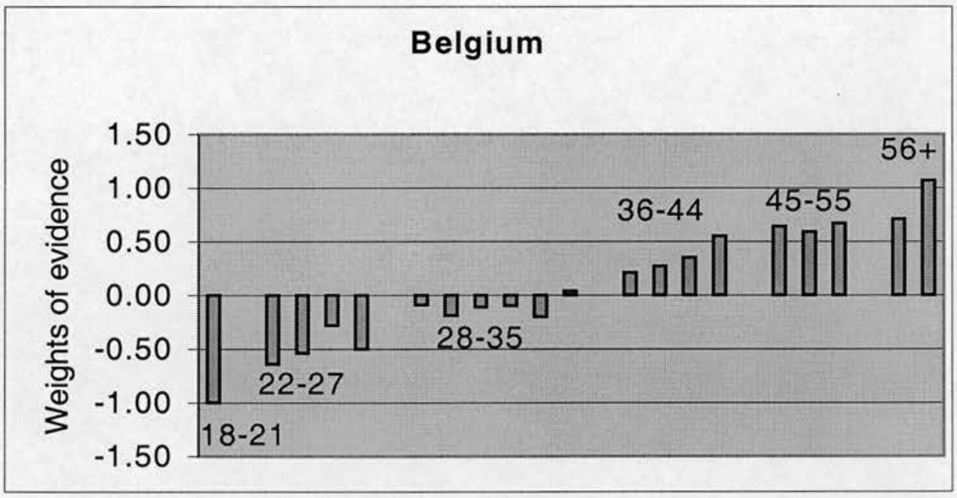
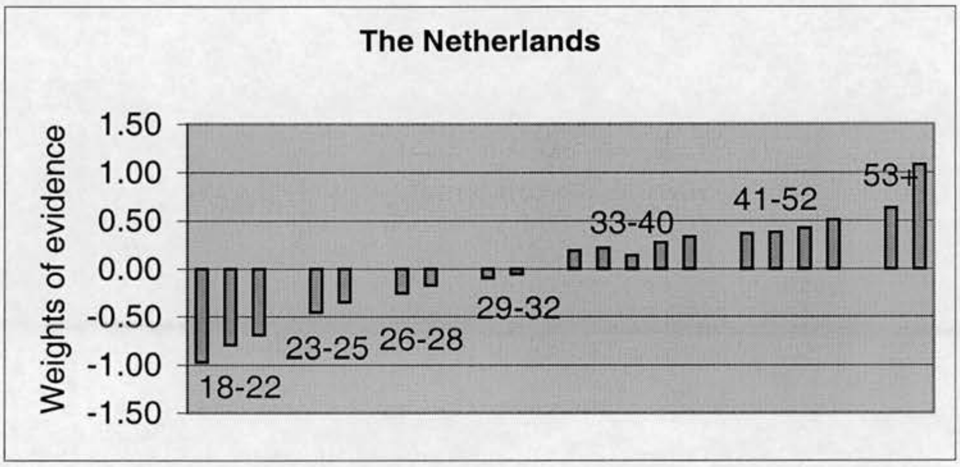
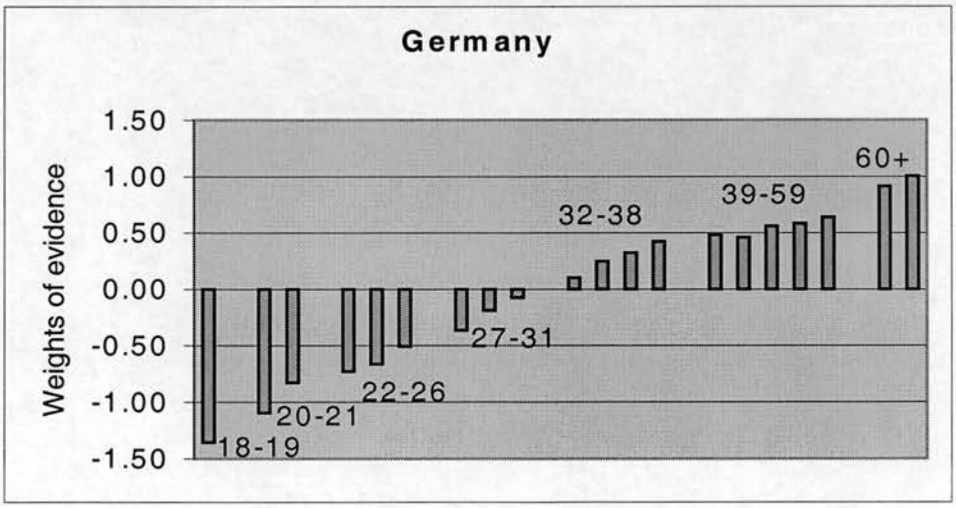


Figure 4.5. Coarse-classification of 'Age' by country. 5% groups.







At the same time although the weights of evidence for the majority of attributes did differ across the countries, it was possible to observe some general patterns that would hold irrespective of the country. For example, for all three countries married and widowed applicants were the best categories, single applicants – the worst, and divorced / living together appeared in the middle. In terms of age, older applicants were better credit risks than young people, although in Germany the youngest age group is relatively riskier than in the other two countries.

Longer times at address and in employment were also associated with better repayment behaviour. Homeowners showed less delinquencies than people in rented accommodation. Those that did not give their home telephone numbers were more likely to default than applicants who indicated their telephones. Self-employed borrowers seemed to be quite a risky category in all three countries, in contrast to retired people and housewives.

Nevertheless, at this stage it was difficult to decide whether the observed similarities were strong enough to compensate for the observed differences.

4.3.3 Performance of national models

The following model was fitted to each of the three national datasets:

$$\log[\frac{\boldsymbol{p}_{ij}}{1-\boldsymbol{p}_{ij}}] = \alpha_j + \beta_{j1}\boldsymbol{x}_{ij1} + \beta_{j2}\boldsymbol{x}_{ij2} + ... + \beta_{jk}\boldsymbol{x}_{ijk}$$

where $\boldsymbol{p}_{ij}$ is the probability of being good for i individual, j country and k predictor variables.

Several logistic regression models were developed for each country, in order then to select the one with the best predictive performance. First of all, two different approaches to variable coding were tried: dummy binary variables and weights of evidence. The former converts all n coarse-classes or bands of the characteristics into n-1 dummy binary variables. The latter assigns some numerical value to each coarse-class, in this case – weights of evidence that were calculated as described in the previous section.

The advantage of binary coding comes from the fact that the resulting coefficient estimates are free from any relationship apart from the one that comes from the estimation algorithm. But this approach leads to a large number of variables. The weights of evidence (WOE) approach reduces the number of variables in the model by giving the attributes an ordering related to the odds of goods to bads in the development sample. But WOE give a value to each attribute which depends only on that characteristic, they fail to account for relations between characteristics.

Tables 4.8 and 4.9 summarise the classification accuracy of the models, which was measured by the area under the ROC curve and the percentage of incorrectly classified observations, when the models were applied to the randomly selected hold-out samples. Models with binary coding slightly outperformed the WOE models for all three countries. Therefore only binary coding was used in the subsequent stages of analysis. The details of binary variable coding for the three countries and the generic model are given in the Appendix, A17 - A20.

The straightforward way to compare the models across the countries would be to force the same characteristics into the model. That was done using binary coding. The predictive performance of the models differed across the countries in spite of the fact that models contained the same characteristics. Judging by the area under the ROC curve (Table 4.8. row 2) the set of common characteristics predicted best of all for the Netherlands, slightly worse for Germany, and worst of all for Belgium. The error rate (Table 4.9) cannot be used for direct comparison between countries, since it depends on a chosen acceptance rate. The rate was fixed to equal the proportion of actual goods in each of the national datasets. Since the Netherlands have slightly lower good rate, it has a lower cut-off compared to Belgium and Germany. However, the error rate can be used for comparing different models within one country.

It would be interesting to compare the coefficients for the same characteristics across the countries, but checks for multicollinearity (tolerance/VIF) revealed that the coefficients in the models were unstable. VIF above 3.6 was considered to signal 'serious' collinearity. Whilst it was not relevant for prediction, it meant that the straightforward comparison of model coefficients was not possible.

Table 4.8. Area under the ROC curve (hold-out sample)

Model type	Variable coding	Selection type	Belgium	The Netherlands	Germany
Main effects	Binary	Stepwise	0.7074	0.7804	0.7394
		Forced	0.7129	0.7814	0.7417
	Weights of evidence	Stepwise	0.6951	0.7625	0.733
Main effects + all interactions	Weights of evidence	Stepwise	0.6979	0.7719	0.7379
Main effects + selected interactions (1)	Binary	Stepwise	0.7050	0.7796	0.7394
Main effects + selected interactions (2)	Binary	Stepwise	0.6929	0.7807	0.7372

Table 4.9. Percent of incorrectly classified observations (hold out sample)

Model type	Variable coding	Selection type	Belgium	The Netherlands	Germany
Main effects	Binary	Stepwise	16.96%	16.50%	15.74%
		Forced	16.92%	16.50%	15.74%
	Weights of evidence	Stepwise	17.38%	17.50%	16.08%
Main effects + all interactions	Weights of evidence	Stepwise	17.32%	16.56%	15.86%
Main effects + selected interactions (1)	Binary	Stepwise	16.92%	16.48%	15.74%
Main effects + selected interactions (2)	Binary	Stepwise	17.40%	16.50%	15.72%

In order to reduce collinearity and develop models with meaningful coefficients, the stepwise procedure of variable selection was applied both for binary and WOE coding. The stepwise selection procedure adds variables to the model one at each step if those variables meet the specified level of significance and at the same time removes variables from the model if they fail to meet the specified level of significance for staying in the model. The level of significance for entry into the model and remaining there was set at 0.05.

Whilst the coefficients in binary stepwise models for Belgium and Germany showed acceptable VIF, the model for the Netherlands still showed significant collinearity. Model respecification was chosen as a method of coping with collinearity. Two approaches were tried. First, in order to account for the possible relationships between the predictor variables, two-level interactions were included into the models. With the WOE coding, adding interactions was not a problem. All possible combinations of products of variables were generated, and the stepwise logistic regression selected the most predictive ones.

However, the same approach with binary coding would have led to a huge number of variables (e.g. 1240 for Belgium) and the logistic regression may not have converged. The number of variables for interactions was reduced in two ways (marked as interactions (1) and interactions (2) in the Tables 4.8 and 4.9). Interactions (1) were selected on the basis of bivariate analysis of association, Cramer's V statistic was chosen as a measure of the strength of association:

$$\boldsymbol{V} = \sqrt{\frac{\chi^2}{(\boldsymbol{N})\boldsymbol{Min}(\boldsymbol{r}-1, \boldsymbol{c}-1)}}$$

Interactions (2) were restricted to cover only the variables selected by stepwise procedure in 'main effects/ binary' model.

So for each country there are six models:

1- Main effects/ Binary/ Stepwise
2- Main effects/ Binary/ Forced
3- Main effects/Weights of evidence/Stepwise
4- Interactions/ Weights of evidence/Stepwise
5- Interactions (1) / Binary / Stepwise
6- Interactions (2) / Binary /Stepwise.

Judging by the area under the ROC curve the best performing models for all three countries were 'Main Effects/Binary/Forced'. Error rate favoured the 'Main Effects/Binary/Forced' for Belgium, 'Interactions(1)' for the Netherlands, and 'Interactions (2)' for Germany. However, it is believed that in the context of the current analysis the ROC-curve is a preferred measure, since it is independent of the cut-off level.

The models giving the best prediction were 'Main Effects/Binary/Forced'. However, as mentioned before, it would be desirable to have models with meaningful coefficients plus the fact that they can predict well. The next best predicting models with coefficients showing acceptable VIF were 'Main Effects/Binary/Stepwise' for Belgium and Germany, and 'Interactions (1)' for the Netherlands (Parameter estimates for these models are given in A21 and VIF is given in A22 of the Appendix).

The tests for significance of differences between areas under ROC-curves as described in section 1.3 conducted at 0.05 level (Table 4.10) were not significant for 'Forced' and 'Stepwise' models for Belgium, and for 'Forced' and 'Interactions (1)' for the Netherlands. But for Germany the difference between 'Forced' and 'Stepwise' was significant at the 0.05 level. Still it was decided to sacrifice a little bit of predictive accuracy in favour of the possibility to interpret the coefficients.

Table 4.10. Tests of significance for AUROC

	Belgium	**The Netherlands**	**Germany**
	Main effects/ Binary/ Forced	Main effects/ Binary/ Forced	Main effects/ Binary/ Forced
Main effects/ Binary/ Stepwise	Not significant	**Significant**	**Significant**
Main effects/Weights of evidence	**Significant**	**Significant**	**Significant**
Interactions/ Weights of evidence	Not significant	**Significant**	**Significant**
Interactions (1)/ Binary	Not significant	Not significant	**Significant**
Interactions (2)/ Binary	**Significant**	Not significant	**Significant**

So the models for comparison between the countries were 'Main Effects/Binary/Stepwise' for Belgium, and 'Interactions (1) / Binary' for the Netherlands. For Germany 'Main Effects/Binary/Stepwise' and Interactions (1) / Binary' showed equally good performance. Since 'Main Effects/Binary/Stepwise' was a more parsimonious model, it was preferred to the model with interactions.

Table 4.11 presents the characteristics in the 'best' model for each country ranked in the order of significance. The rankings were obtained by calculating the distance between the highest parameter and the lowest one (including the reference group) of binary variables representing categories of one characteristic. E.g, from Table A21 in the Appendix for Belgium the highest parameter for age is 0.21 (AGE5), the lowest is –0.28 (AGE1). The distance of 0.49 is used as measure of separation between categories of AGE. The greater distance is interpreted as an indication of better discriminating power, and therefore, higher significance.

The models do look very different, e.g. 'Marital status' is not in the Belgian model at all, and for the variables that appear in all three models, the ranks are not the same. Although 'Business type' is the most predictive characteristic for both Germany and the Netherlands.

Table 4.11. Ranks of variables in national models

Rank	Belgium	The Netherlands	Germany
1	Occupation	Business type	Business type
2	Number of dependants	Telephone	Occupation
3	Time at address	Goods code*Payment date	Spouse age
4	Residential status	Marital status	Applicant's age
5	Goods code	Applicant's age	Telephone
6	Payment date	Occupation	Time on job
7	Goods price	Residential status	Goods code
8	Telephone	Number of dependants	Number of dependants
9	Time on job	Time at address	Time at address
10	Spouse age	Credit insurance	Card insurance
11	Business type	Goods price	Residential status
12	Applicant's age	Time on job	Payment date
13	Credit insurance	Spouse age	Marital status
14	Employer's phone		

These models differ between the countries:

1) in attributes for each characteristic and in coarse-classing the attributes;
2) in characteristics selected by the stepwise routine;
3) different WOE and β-values for any common attributes or coarse-classes.

So it is possible to conclude, judging on univariate analysis and regression analysis, that changes in population (population drift, see Section 3.2) occur in these samples between the countries. These changes manifest through differences in:

1) in class (good/bad) priors, $p(i)$ (Table 4.6),
2) in the posterior distributions of class membership $p(i|x)$ (Table 4.9).

4.4 Generic model vs national models

4.4.1 Model specification and predictive performance

In this section the predictive performance of a model which is generic to all three countries is compared with the performance of the 'best' national models. The definition of good/bad used in the generic model was the same as in the country-specific analysis. Three national samples were used for the development and validation of the generic model. The whole of the Belgian sample was taken, since it was the smallest one. Proportional stratified samples were taken from the Dutch and German samples to keep roughly the same proportions of 'good', 'bad' and rejected applicants in the subsamples as in each of the samples. Although proportional stratification is not required when the logistic regression is used for modelling (Hosmer and Lemeshow (2000)), it was carried out so that cut-offs for the confusion matrix were not distorted by the subsampling process.

The number of accepted applicants was chosen to be almost the same as in the Belgian dataset, so that no country would dominate in the analysis. The frequency of classes by country in the resulting subsamples is given in Table 4.12. The numbers in brackets show the percent of the category in the overall generic sample. The proportions of good/bad/rejected and accepted are not exactly as desired, since it was impossible to make the random sampling procedure to return exactly the same proportions. The aggregated dataset was randomly split into a training set (70%) and a hold-out validation sample.

Table 4.12. Frequency of goods/bads/rejected by country in the aggregated dataset

	Belgium	Germany	The Netherlands	Total
Good	23200 (17.31%)	23204 (17.31%)	22785 (17.00%)	69189 (51.63%)
Bad	3090 (2.31%)	3097 (2.31%)	3620 (2.70%)	9807 (7.32%)
Rejected	17966 (13.41%)	21290 (15.89%)	15769 (11.77%)	55025 (41.06%)
Total	44256 (33.02%)	47591 (35.51%)	42174 (31.47%)	134021 (100%)

The characteristics were coarse-classified according to the procedure described in Section 4.3.2 but without any reference to the country. The generic model was first estimated without any variables to indicate which country the case was a member of, and subsequently, the country indicator variables were included into the analysis. Based on the arguments of Chapter 2, one cannot say definitively whether the use of such indicator variables is legally acceptable. They were included into the analysis in order to test if the country indicators were statistically significant and if they had any impact on the predictive ability. Although the country indicators demonstrated significance at 5% level (see A21 in Appendix), their impact on the predictive accuracy was minor (Tables 4.13-4.14). It should be noted though that the inclusion of these variables changes the relative ranking of applicants from different countries and therefore changes acceptance rates for different nationalities.

The predictive ability of the model was measured by the area under the ROC curve, when applying the model to the hold-out sample of the aggregated dataset and calculating the percentage of incorrectly classified applicants. The cut-off was chosen so that the actual number of bads equalled the predicted number of bads in the hold-out sample. The generic model was also applied to each of the national hold-out samples separately in order to compare the predictive performance of the generic model to each of the national models.

Table 4.13 Generic model performance -Stepwise. AUROC and error rate

Hold-out sample	Model type	Belgium	The Netherlands	Germany
AUROC				
Generic	Generic - no 'country' dummy variable	0.745		
	Generic - 'country' dummy variable included	0.746		
National	Generic - no 'country' dummy variable	0.701	0.777	0.731
	Generic - 'country' dummy variable included	0.705	0.778	0.729
	National	0.707	0.780	0.739
ERROR RATE				
Generic	Generic - no 'country' dummy variable	16.54%		
	Generic - 'country' dummy variable included	16.72%		
National	Generic - no 'country' dummy variable	17.00%	16.58%	16.04%
	Generic - 'country' dummy variable included	16.92%	16.56%	16.24%
	National	16.96%	16.48%	15.74%

The rationale for selecting the model for comparison was the same as in country-specific analysis: select the models that give the best prediction and at the same time with coefficients not inflated by collinearity. Table 4.13 presents the area under the ROC-curve and error rate for national and generic models that satisfy the conditions given above. The variables with parameter estimates are given in A21 in the Appendix. Table 4.14 presents the performance of models with all binary variables forced to enter. The difference in prediction between models with stepwise selection and models with all variables is minimal.

The results demonstrate that, in general, the performance of the generic model is very close and comparable to that of the national ones. The tests on national hold-out samples do favour the national models slightly, but the difference is marginal. In fact, significance tests at 0.05 level (Table 4.18) show that there is a significant difference between generic and national models only for Germany. For Belgium and the Netherlands the difference is not significant at 0.05 level.

Table 4.14 Generic model performance- Forced. AUROC and error rate

Hold-out sample	Model type	Belgium	The Netherlands	Germany
AUROC				
Generic	Generic - no 'country' dummy variable	0.746		
	Generic - 'country' dummy variable included	0.748		
National	Generic - no 'country' dummy variable	0.703	0.779	0.732
	Generic - 'country' dummy variable included	0.705	0.780	0.732
	National	0.713	0.781	0.742
ERROR RATE				
Generic	Generic - no 'country' dummy variable	16.48		
	Generic - 'country' dummy variable included	16.58		
National	Generic - no 'country' dummy variable	17.02	16.68	16.16
	Generic - 'country' dummy variable included	16.90	16.60	16.16
	National	16.92	16.50	15.74

4.4.2. Differences between applications accepted by different models

Although the percentage of applications classified incorrectly by different models was roughly the same, meaning that there might be little difference for the lenders as to which model to use, the rejected applications were different in each case.

For each country hold-out sample the proportions of applicants accepted by the generic model but rejected by the national model and vice versa were calculated. For Belgium these proportions constituted 3.83% each from the total hold-out sample, for the Netherlands 2.75%, and for Germany 2.82%. In other words, the decision whether to grant credit or not would be identical (irrespective of whether the generic or national decision rule is used) for 92.34% of Belgian applicants, 94.50% of the Dutch, and 94.36% of German applicants.

However, it would be of interest to investigate what characteristics distinguish between the two groups:

1. the applicants predicted 'Good' (and therefore accepted) by the generic model and predicted 'Bad' (and rejected) by the country model;
2. the applicants predicted 'Good' by the country model and 'Bad' by the generic model.

So for each country the two groups were cross-examined by looking at frequency distributions for categorical characteristics, and means and medians[10] for continuous characteristics.

Tables 4.15 - 4.17 present the characteristics that showed the most striking differences. For categorical characteristics only some attributes are selected for the purpose of illustration, that is why the percentages within one characteristic do not always sum to 100%. As a benchmark, the frequency distributions, means and medians are also reported for the total hold-out sample.

For all countries applicants accepted by the generic model, but rejected by the national one would be slightly older than those accepted by the national model, but rejected by the generic one: a median age for former would be 29 years for both Belgium and the Netherlands versus 25 and 23 respectively. For Germany the corresponding median ages would be 25 and 22 years old.

In Belgium those applicants accepted by the generic model, but rejected by the national one, would have a significantly higher percentage of those working in the industry, divorced and renting a flat, while the percentage of those working in the service sector would be much lower. Applicants with the card insurance would clearly prefer to be scored by the national model rather than by the generic one, since the former accepted more applicants with the insurance.

In the Netherlands the generic model accepted a higher percentage of officials, those using their cards for 'card applications' and widowed, as compared to the national model, that accepted higher percentages of applicants working in catering, construction and those buying TV-sets.

[10] Most distributions were skewed due to high incidence of 0 values

Table 4.15a). Differences between the applicants accepted by one model but rejected by another one. Belgium: categorical characteristics

Characteristic	Attribute	Total	Accepted by generic, rejected by country (3.83%)	Accepted by country, rejected by generic (3.83%)
		% cases		
Business type	Industry	18.56%	28.24%	15.61%
	Service prof	7.54%	2.99%	14.29%
Card Insurance	No Insurance	82.77%	85.71%	73.09%
	Insurance	17.18%	14.29%	26.91%
Goods code	Phones	11.54%	17.61%	29.90%
Marital Status	Single	27.09%	45.18%	64.12%
	Divorced	12.54%	16.61%	6.98%
Occupation	Part-Time	5.10%	8.31%	2.66%
	Self-Employed	3.96%	6.64%	11.96%
Telephone	Given	89.01%	92.03%	73.09%
Employer's phone	Not given	45.75%	37.54%	48.17%
	Given	54.20%	62.46%	51.83%
Residential Status	Rented Flat	18.70%	28.90%	16.61%
	Living with parents	12.94%	26.25%	35.55%

Table 4.15b). Differences between the applicants accepted by one model but rejected by another one. Belgium : continuous characteristics

	Mean			Median		
	Total	**Accepted by generic, rejected by country**	**Accepted by country, rejected by generic**	**Total**	**Accepted by generic, rejected by country**	**Accepted by country, rejected by generic**
Age	38	31	28	37	29	25
Spouse age	38	30	30	37	29	27
Number of children	0.89	0.65	0.69	0	0	0
Time at address	7yr 9m	2yr 2m	4yr 5m	7yr 1m	2yr 10m	6yr 8m
Time on job	7yr 9m	2yr 3m	2yr 7 m	5yr 8m	1yr 6m	1yr 8m
Goods price (Euro)	646.93	835.46	485.24	421.17	818.00	507.77

Table 4.16a). Differences between the applicants accepted by one model but rejected by another one. The Netherlands: categorical characteristics

Characteristic	Attribute	Total	Accepted by generic, rejected by country (2.75%)	Accepted by country, rejected by generic (2.75%)
		% cases		
Business type	Catering	3.16%	1.79%	11.76%
	Construction	2.38%	0.45%	7.14%
	Officials	3.52%	5.80%	0.15%
Credit Insurance	No Insurance	70.50%	72.02%	61.61%
Goods code	Card Applications	6.71%	20.98%	0.60%
	HIFI Radio	5.94%	4.17%	23.21%
	TV	8.78%	2.23%	14.14%
	Mopeds	3.21%	1.19%	12.50%
Marital Status	Single	33.97%	59.52%	65.03%
	Widowed	11.65%	11.46%	6.25%
Occupation	Benefit	16.68%	11.76%	15.92%
Telephone	Mobile	7.36%	29.76%	14.58%
	Given	87.98%	58.93%	73.21%
Residential Status	Living with parents	8.88%	18.60%	26.79%

Table 4.16b). Differences between the applicants accepted by one model but rejected by another one. The Netherlands: continuous characteristics

	Mean			Median		
	Total	**Accepted by generic, rejected by country**	**Accepted by country, rejected by generic**	**Total**	**Accepted by generic, rejected by country**	**Accepted by country, rejected by generic**
Age	37	31	30	36	29	23
Spouse age	38	34	38	37	32	36
No of children	0.74	0.13	1.17	0	0	1
Time at address	5yr 10m	4yr 7m	2yr 1m	3yr 7m	2yr	1yr
Time on job	7yr 1m	5yr	2yr 11m	4yr 6m	2yr 5m	1yr 2m
Goods price (Euro)	903.60	571.23	916.62	687.25	308.12	793.21

Table 4.17a). Differences between the applicants accepted by one model but rejected by another one. Germany: categorical characteristics

Characteristic	Attribute	Total	Accepted by generic, rejected by country (2.82%)	Accepted by country, rejected by generic (2.82%)
		% cases		
Business type	Services	29.31%	30.50%	40.00%
	Construction	14.47%	8.33%	6.92%
	Shop Employee	6.76%	10.22%	6.15%
Card Insurance	No Insurance	90.48%	92.92%	73.11%
	Insurance	9.52%	7.08%	26.89%
Goods code	HIFI Radio	13.35%	33.18%	18.87%
Marital Status	Married	52.36%	17.92%	12.42%
	Single	32.96%	75.31%	77.04%
Occupation	Full-Time	82.72%	82.23%	91.82%
	Self-Employed	4.15%	12.74%	2.83%
Telephone	Not given	5.76%	12.58%	15.72%
Residential Status	Living with parents	11.88%	22.96%	39.78%

Table 4.17b). Differences between the applicants accepted by one model but rejected by another one. Germany: continuous characteristics

	Mean			Median		
	Total	Accepted by generic, rejected by country	Accepted by country, rejected by generic	Total	Accepted by generic, rejected by country	Accepted by country, rejected by generic
Age	37	28	25	36	25	22
Spouse age	17	13	12	0	0	18
Number of children	0.72	0.47	0.22	0	0	0
Time at address	6yr 9m	2yr 9m	1yr 8m	4yr	2yr 6m	1yr 9m
Time on job	5yr 10 m	2yr 7m	1yr 4m	4yr 4m	2yr 3m	1y 3m
Goods price (Euro)	794.45	761.09	554.22	715.30	715.81	459.65

The German applicants favoured by the generic model, but rejected by the country-specific model would have higher percentage of shop employees, self-employed and married. On the contrary, the national model accepted more people with card insurance, in full-time employment and living with parents.

So for the consumer it does make a difference which model is applied.

4.4.3. Effect of incorporating additional information

The comparison of generic models with national models would be incomplete without incorporating full information that was available for each country into the analysis. Therefore, a third type of model was built – national models were developed on a full set of characteristics, both application and CRA.

Binary coding was used as in the case of national and generic models developed on a set of common characteristics, and a stepwise selection procedure was applied in order to get the most predictive variables. The performance of 'full information' models was compared to 'best' national and generic (no 'country' indicator variable) models and is summarised in Table 4.19 and Figure 4.6.

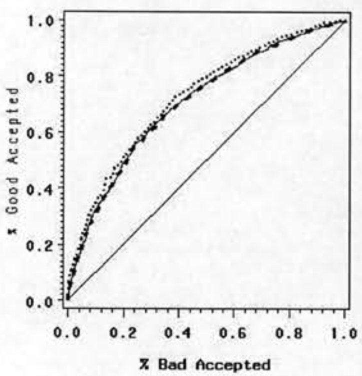
Tests of significance at 0.05 level (Table 4.18) showed that for Belgium and the Netherlands there was a significant difference between the 'full information' and the other two models, but not between the generic and national models developed on a set of common characteristics. For Germany all three models were significantly different.

Table 4.18. Generic model performance. Significance tests on AUROC, 0.05 level.

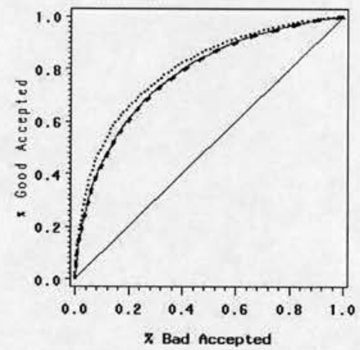
	National models		
	Belgium	**The Netherlands**	**Germany**
Generic model, no country indicator	Not significant	Not significant	Significant
'Full information' models	Significant	Significant	Significant

Figure 4.6. ROC-curves for three types of models by country

Belgian hold–out



Dutch hold–out



German hold–out

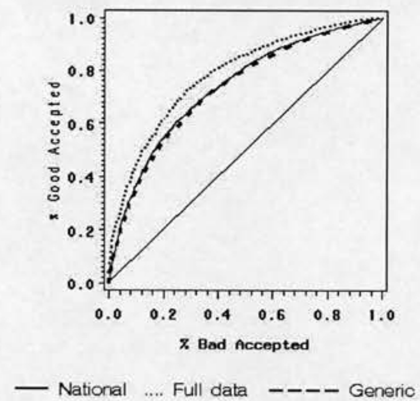


Table 4.19a) Three types of models. AUROC

	Belgium	The Netherlands	Germany
Generic	0.701	0.777	0.731
National	0.707	0.780	0.739
Full-info	0.722	0.800	0.789

Table 4.19b) Three types of models. Error rate

	Belgium	The Netherlands	Germany
Generic	17.00%	16.58%	16.04%
National	16.96%	16.48%	15.74%
Full-info	16.72%	15.94%	14.68%

The new variables that entered the model are summarised in Table 4.20, more details on 'full information' models are given in Appendix, A23.

Table 4.20 Additional characteristics entering 'full information' models

	Belgium	**The Netherlands**	**Germany**
1	*Credit card type*	Negative CRA record	*Credit card type*
2	*Time at bank*	Number of live fixed term accounts	*Time at bank*
3	Previous employment Given/ Not given	Number of live revolving accounts	Total amount of live credit
4	*Initial instalment*	Instalment paid	*Initial instalment*
5	Retailer	*Dealer*	*Section/ Dealer*
6	Spouse's occupation	Loan amount	Spouse marker Yes/No
7	*Type of agreement*	*Type of agreement*	*Type of agreement*
8	Total addresses given	Time since last bad 'A' account	Time at CRA 1
9		Time since last bad 'A+' account	Time at CRA 2
10		*Time first CRA registration*	*Time at CRA*
11		*Deposit*	*Deposit*
12		Bank code	Bank sort code
13		*Nationality*	*Nationality*
14			Loan to instalment
15			Theoretical instalment
16			Percent deposit
17			Underwriter
18			East/ West Germany indicator
19			Worst CRA credit
20			CRA quantity 1
21			CRA quantity 3
22			CRA quantity 4
23			Type of bank account

Overall, it is possible to conclude that incorporating additional information into the models does increase their predictive ability, which is especially evident for Germany, where 23 new variables, including 8 CRA characteristics, entered into the model. The improvement compared to national models ranges from 0.015 (Belgium) to 0.050 (Germany) in terms of AUROC, and as for error rate the improvement ranges from 0.24% (Belgium) to 1.06% (Germany).

It is notable that quite a few characteristics (marked in italics in Table 4.20) are available for 2 countries. Therefore, there is a significant potential for expansion for generic models, provided the data collection practices are harmonised across the countries.

'Nationality' turned out to be significant for Germany and the Netherlands. Whilst the inclusion of the variable was appropriate at the time when the analysis was done, the implementation of the Race Equality Directive (Council of the EU (2000)) changed the situation. In the Netherlands 'nationality' is interpreted as being part of 'race', so the use of this information can be regarded as discrimination (Koopman (1999), Zwamborn (2003)). In Germany the Directive is not transposed into the national legislation yet, so the current legal status of nationality is not clear (Mahlman (2003)).

It is obvious that the difference between generic and national models built on the same list of characteristics is less pronounced. It can be attributed to the flat maximum effect, which can compensate for differences in grouping and weighting of similar characteristics. The difference between generic and national models ranges from 0.003 (the Netherlands) to 0.008 (Germany) for AUROC, and from 0.04% (Belgium) to 0.3% (Germany) for error rate. But even in the case of Germany where the gap between the models is largest, it can be argued that the difference is only marginal. Thus generic scoring is potentially a viable option.

4.5 Conclusions

This Chapter presented an investigation of risk patterns in three European countries and compared the performance of the generic model to that of models built for each country separately. The combination of countries analysed in this Chapter was not explored in previous studies.

In contrast to previous findings, the generic model built for different populations of three European countries showed an adequate performance comparable to that of the national models. This can be attributed to the relative similarity of the countries used in the analysis. This meant that it was possible to select a set of characteristics that could be harmonised across three countries. This set contained enough information to allow for good classification. Although there were differences in distributions of good/bad classes across the countries, there were general patterns that could be observed and the differences were compensated by the flat maximum effect. In addition, a generic model was developed on a heterogeneous sample, in which all three populations were appropriately represented.

Nevertheless, whilst the generic model showed an acceptable level of predictive performance, the applicants accepted with the generic decision-rule differed from those accepted with the country-specific rule on a number of characteristics.

This gives some support to the argument that legislative restrictions on data used in credit scoring would harm consumers. Whereas certain characteristics (in this case country of residence) are not available, the lenders may still achieve the same level of classification accuracy, but the absence of the relevant information may have a differential and/or adversarial effect for certain groups of applicants.

Additional information increases the predictive ability of models, whilst the difference between generic and national models developed on the same set of common characteristics is marginal. This emphasises the value of information and the need for harmonisation of data across Europe. One may expect that further European harmonisation will increase the potential scope of application of generic models.

Chapter 5. European generic scoring using survival analysis

5.1. Introduction

Traditionally credit scoring has been concerned with the estimation of the probability of default, i.e. the level of risk presented by the credit applicant. However, it has been shown by Hopper and Lewis (1992) and Leonard (1997) that the level of risk is not necessarily a good indicator of the level of profit the applicant will generate. Some high risk applicants can generate a significant profit if they use the credit product actively and pay interest and charges long enough before going into default. On the contrary, low risk applicants may pay the full balance every month, thus keeping the revenues from such 'good' accounts low. These revenues may not be enough to cover the costs associated with maintaining the accounts.

In recent years the focus of credit risk modelling shifted from risk to profit (Thomas et al. (2002)). In terms of profitability of an account, one of the most important aspects is its lifetime, i.e. how long the credit applicant uses the account. Modelling lifetime is the domain of survival analysis. Whilst survival analysis is a well-established technique in medical research, reliability and engineering (Collett (1994), Ansell and Phillips (1994)) , there have been only a few applications in credit scoring (Narain (1992), Banasik et al. (1999), Stepanova and Thomas (2001), Hand and Kelly (2001), Till (2001), Baesens (2003)). When compared to logistic regression, survival analysis methods showed superior performance in certain applications in credit scoring (Narain (1992), Hand and Kelly (2001))and inferior performance in some other applications, depending on the length of time used to measure the classification accuracy and the nature of the problem being modelled (Banasik et al. (1999), Stepanova and Thomas (2001), Baesens (2003)).

The studies conducted so far, investigated fixed term loans and populations coming from one country. In this Chapter the data on a revolving type of credit is analysed with the aim to predict the time to default. Continuing the work started by previous research projects that focused on some specific approach within survival analysis, this Chapter investigates a range of distributions for accelerated lifetime / failure time models and a semi-parametric proportional hazards model. These are benchmarked against logistic regression.

Apart from the comparison of different approaches within the survival analysis, there is a second objective: to investigate the sensitivity of the predictive ability of survival models, as compared to logistic regression, to the presence of heterogeneous subpopulations, i.e. applicants from three different European countries.

The Chapter is structured in the following way. Section 5.2 presents an overview of the basic concepts and methods used in lifetime modelling. Section 5.3 surveys previous applications of survival analysis to credit scoring problems. Section 5.4 compares the national survival patterns and estimates the difference in predictive accuracy of national models versus the generic model. It also investigates the performance of generic models that incorporate stratification and time-dependency. Section 5.5 concludes.

It is concluded that for the dataset used in this thesis, survival analysis provides at least as good classification accuracy as logistic regression, whilst offering some additional benefits. Different survival analysis methods are very close to each other in predictive performance. And the difference in predictive accuracy of generic and national survival models is marginal, in line with the results from Chapter 4.

5.2. Survival analysis concepts and methods

5.2.1. Describing lifetime distributions

In the previous chapters the analysis focused on estimating the probability of the event occurring, the event being defined as '2 consecutive months in arrears within the observation period (12-25 months)'. Survival analysis estimates the time to an event, which is termed as a lifetime or survival time, and is assumed to be a realisation of some random process. There are four standard ways to describe T, the event time:

1. Cumulative distribution function (c.d.f.), which indicates the probability that the event will occur before any time, t, chosen (including t):

$F(t) = Pr\{T \leq t\}$.

It is more common to use the related function, the survivor function

$S(t) = Pr\{T > t\} = 1 - F(t)$,

which gives the probability of the event not occurring before the specified time t or the probability of surviving beyond t.

2. Probability density function, p.d.f.

$$f(t) = P\{t \le T \le t + \Delta t\} = \lim_{\Delta t \to 0} \frac{P\{t \le T \le t + \Delta t\}}{\Delta t} \qquad (5.1)$$

which gives the probability of event occurring within the time interval $(t, t+\Delta t)$. It can be also expressed as

$$f(t) = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt} \qquad (5.2)$$

3. Hazard function

$$h(t) = \lim_{\Delta t \to 0} \frac{P\{t \le T \le t + \Delta t \mid T \ge t\}}{\Delta t} \qquad (5.3)$$

which is the probability of the event occurring within the time interval $(t,t+\Delta t)$, provided the event has not occurred so far. The hazard function is interpreted as an instantaneous failure rate and is often referred to as the conditional density or intensity function. It can be also expressed as:

$$h(t) = \lim_{\Delta t \to 0} \left\{ \frac{F(t + \Delta t) - F(t)}{\Delta t} \right\} \frac{1}{S(t)} = \frac{f(t)}{S(t)} = -\frac{d}{dt}\{\log(S(t))\} \qquad (5.4)$$

3. Cumulative or integrated hazard function

$$H(t) = \int_0^t h(u)du = -\log(S(t)) \qquad (5.5)$$

and alternatively, the survival function can be expressed as

$$S(t) = \exp(-H(t)) \qquad (5.6)$$

All these functions are essentially equivalent ways to describe a lifetime distribution and if one is given, the others can be derived, as shown above.

5.2.2. Censoring and competing risks

One of the advantages of survival analysis is its ability to handle censored cases, i.e. the cases for which the event of interest was never observed. Censoring comes in several forms. Right censoring occurs when T is greater than some value c, which typically happens because the observation period is terminated before the event is observed. Left censoring refers to situations when the process starts prior to the start of the observation so that the latter part of the lifetime is seen. The combination of both right and left censoring is defined as interval censoring.

Within right censored cases it is possible to distinguish Type I, Type II and random censoring. Type I occurs when the observation period is terminated at a particular point in time. Type II occurs when the observation period is terminated after a prespecified number of events have occurred. Random censoring occurs when the observation period is terminated due to the reasons beyond the control of the researcher. With random censoring, it is important to be sure that it is non-informative, i.e. that the censored time is independent of the unobserved event time, otherwise, it can lead to severe biases.

Another important advantage of survival analysis is that it allows the researcher to discriminate between different reasons for an event. This approach is called the competing risks. In application to revolving credit one may want to predict not only the default, but also the possibility of closing an account too early, e.g. the instances when only one purchase is made on the card, and then the card is never used again.

In this case the analysis is done separately for each event type. When predicting default, the accounts that have not experienced default, are considered censored. When predicting early closure, all the open accounts, including the delinquent ones, are considered censored.

However, the competing risk approach is not pursued in this thesis. It will be shown in Section 5.4.1 that there are some accounts that have been recorded to be closed before the end of the observation period. It will also be shown in Chapter 6 that an open account does not necessarily imply it is being used. So 'early closure' in the context of the data being analysed is not straightforward to define. Chapter 6 will discuss this issue in more detail.

5.2.3. Some common lifetime distributions

There are several distributions of survival times that are frequently used in medical studies (Collett (1994)) and engineering (Ansell and Phillips (1994)).

The *exponential* model assumes that the hazard is constant over time: $h(t) = \lambda$; the survivor function is then $S(t)=e^{-\lambda t}$, and the implied p.d.f. of survival times is $f(t) = \lambda e^{-\lambda t}$ for $0 \leq t < \infty$.

An important property of the exponential distribution is the lack of memory, i.e. the distribution of additional survival times is not affected by the knowledge that an individual has survived for a certain length of time.

The *Weibull* model allows the hazard rate to decrease or increase monotonically with time.

$h(t) = \lambda \gamma t^{\gamma-1}$

$S(t) = \exp(-\lambda t^{\gamma})$

$f(t) = \lambda \gamma t^{\gamma-1} \exp(-\lambda t^{\gamma})$ for $0 \leq t < \infty$.

The shape of the hazard function depends on γ, which is known as a shape parameter, while λ is a scale parameter. When $\gamma=1$, the distribution becomes exponential.

The *log-normal* model implies a non-monotonic hazard function. It rises from 0 to a peak then declines towards 0 as t goes to infinity. Survival time T is said to have a log-normal distribution with parameters μ and σ if $\log T$ has a normal distribution with mean μ and variance σ^2. The p.d.f. of T is given by

$$f(t) = \frac{1}{\sigma\sqrt{(2\pi)}} t^{-1} \exp\left(\frac{-(\log t - \mu)^2}{2\sigma^2}\right) \qquad \text{for } 0 \leq t < \infty.$$

The survivor and hazard functions can be expressed in terms of incomplete normal integrals (Kalbfleisch and Prentice (1980) provide details). In the presence of censoring, estimation becomes complicated, so very often the log-logistic model is used instead of the log-normal. The log-normal is often appropriate for modelling repeatable events.

The *log-logistic* distribution can be used as an approximation to log-normal to model a non-monotonic hazard function with a single mode:

$$h(t) = \frac{e^{\theta} kt^{k-1}}{1 + e^{\theta} t^{k-1}}$$

$$S(t) = [1 + e^{\theta} t^{k}]^{-1}$$

$$f(t) = \frac{e^{\theta} kt^{k-1}}{(1 + e^{\theta} t^{k})^{2}}$$

All distributions considered so far (apart from log-logistic) are special cases of the generalised *gamma* model, which can take a wide variety of shapes, including the bathtub ones. It can also represent hazard functions that have more than one peak. The flexibility of this model comes at a cost of increased complexity of estimation, and with some convergence problems. The hazard function for the gamma distribution with mean ρ/λ and variance ρ/λ^2 is

$$h(t) = \frac{\lambda^{p} t^{p-1} e^{-\lambda t}}{\Gamma(\rho)\{1 - \Gamma_{\lambda t}(\rho)\}},$$

where $\Gamma(\rho)$ is a gamma function and $\Gamma_{\lambda t}(\rho)$ is an incomplete gamma function given by

$$\Gamma_{\lambda t}(\rho) = \frac{1}{\Gamma(\rho)} \int_0^{\lambda t} u^{\rho-1} e^{-u} du.$$

Collett (1994) states that the Weibull and gamma distributions will normally give very similar results. Still fitting a generalised gamma model may be useful for making comparisons of the goodness-of-fit between different models, since they can be considered as nested within gamma.

5.2.4. Non-parametric estimation of survivor and hazard functions

The most common non-parametric approach to estimating the survivor function, also known as the survival curve, is the Kaplan-Meier (KM) or product-limit estimator.

$$\hat{S}(t_j) = \prod_{j=1}^{i} [1 - \frac{d_j}{n_j}] \tag{5.7}$$

where d_j is the number of events at time j and n_j is the number at risk at time j. The quantity in brackets can be interpreted as the conditional probability of surviving to time t_{j+1}, provided one has survived to time t_j. It was shown by Kaplan and Meier (1958) that the product-limit was in fact, a maximum likelihood estimator.

There are two useful transformations of survival curves that provide information about the shape of the underlying hazard function:

1. A plot of negative log of the survivor function against time, a LS plot, indicates whether the hazard is constant, increasing or decreasing with time. A constant hazard corresponds to a straight line with an origin at 0. A LS plot is in fact, a plot of the cumulative hazard function (5.5). If $h(t)$ is some constant λ, $H(t)=\lambda t$, this correspondents to the exponential distribution.
2. A plot of $log[-log\hat{S}(t)]$ versus $log\ t$, a LLS plot, shows whether survival times decrease or increase monotonically with time. So a straight line with a slope not equal to unity would suggest that a Weibull distribution is suitable for modelling. (Parmar and Machin (1995)).

It is also possible to estimate the hazard function from (5.7), see Stepanova (2001):

$$\hat{h}(t_j) = \frac{d_j}{n_j} = 1 - \frac{\hat{S}(t_j)}{\hat{S}(t_{j-1})} \tag{5.8}$$

Estimating and plotting the survival curves and hazard functions is a powerful tool for initial data exploration, which can provide an indication of the distribution of the hazard function and thus aid in selecting the subsequent methods of analysis.

If there exist sub-populations then it may be desirable to explore whether there are differences in the survivor functions between the sub-populations. The most commonly used tests for the equality of the survival curves are Log-rank and Wilcoxon tests (Kalbfleisch and Prentice (1980)).

The Log-rank test involves calculating deviations of the observed number of events from the expected numbers. For any group in the population the Log-rank statistic is

$$\sum_{j=1}^{r} (d_{ij} - e_{ij})$$

where d_{ij} and e_{ij} are observed and estimated numbers of events occurring in group i at time j, and r is the unique event time in the population. The test of equality of the k survival curves is based on the χ^2_{k-1} distribution for the squared values of the log-rank statistic divided by estimated variance.

The Wilcoxon test is very similar to the procedure described above. The only difference lies in the fact that the Wilcoxon test calculates the weighted sum of deviations of observed numbers of events from expected numbers:

$$\sum_{j=1}^{r} n_j (d_{ij} - e_{ij})$$

where n_j is the total number at risk at each time point.

It follows that the log-rank test is more suitable for detecting proportional differences between k survival curves $S_1(t)=[S_2(t)]^{\gamma}$, where γ is some positive number not equal to 1. The Wilcoxon test is more suitable when event times follow log-normal distributions - it gives more weight to earlier times. However, both methods fail to give reliable results when survival curves cross (Allison (1995)).

5.2.5 Estimating regression models

The methods of the previous section can be used to test for heterogeneity in the population, but with a large number of variables that can potentially influence the survival time, the analysis becomes cumbersome. Regression analysis provides a more comprehensive analysis of the relationship between the survival time and various covariates, so that the predictions could be made about the length of the survival or the probability of survival to a specified time based on the observed covariates.

There are two standard ways to relate time T to predictor variables or covariates: proportional hazards (PH) model and accelerated lifetime/ failure time (AFT) models. The choice of a model will depend on whether it is believed that the covariates act multiplicatively on survival time T, or that they act multiplicatively on the hazard rate h. In the former case an accelerated failure time (AFT) model would be considered suitable, in the latter, a proportional hazards (PH) model. The Weibull family of distributions can be both AFT and PH. However, the PH approach has an advantage in that this model does not require specification of the underlying distribution.

5.2.5.1. Accelerated failure time models

The general form of the AFT models is given by

$$\boldsymbol{h(t) = h_0 (te^{\beta' x})e^{\beta' x}} \tag{5.10}$$

where h_0 are baseline hazard functions, x is the vector of covariates and β is the vector of parameters that need to be estimated. This model specifies that the covariates act multiplicatively on T, so the effect of covariates is to increase or decrease the hazard rate with time. It also implies that the baseline hazard function exists and has some specific distribution.

The AFT models can also be expressed as log-linear models:

$\log \boldsymbol{T} = \boldsymbol{\beta}'x + \sigma\varepsilon$ or $\boldsymbol{y} = \boldsymbol{\beta}'x + \sigma\varepsilon$,

where ε is an error term, and σ is the parameter which changes the distribution of ε, thus acting as a switch between different distributions of T. The relationship between the distributions is as follows:

Distribution of ε	Distribution of T
extreme value (1 parameter)	exponential
extreme value (2 parameters)	Weibull
log-gamma	gamma
logistic	log-logistic
normal	log-normal

Source: Allison (1995).

In the case of no censoring, the log likelihood L can be written as

$L = \Sigma \log(f(w_i)/\sigma))$, where $w_i=(y_i - \beta' x_i)/\sigma$.

With the right-censored observations, the log likelihood becomes

$L = \Sigma \log(f(w_i)/\sigma) + \Sigma \log(S(w_i))$,

where the first sum is over uncensored observations and the second sum is over right-censored observations.

The log-likelihood function is maximised by means of some numerical method. In this thesis a ridge stabilised version of the Newton-Raphson algorithm as implemented in SAS is used.

5.2.5.2. Cox proportional hazards model

In the proportional hazard model the covariates act multiplicatively on the hazard rate:

$h(t)=h_0(t)\exp(\beta' x)$

The baseline hazard function $h_0(t)$ can be regarded as the hazard function when all covariates are equal to 0. Taking logarithms of both sides leads to

$\log h(t)=\alpha(t)+ \beta' x$,

if $\alpha(t)= \alpha$, it is the exponential model;

if $\alpha(t)= \alpha \log t$, it is the Weibull model.

However, the main attraction of the PH model is that there is no need to specify the distribution. Under the assumption of proportionality, the hazard for any individual i is a fixed proportion of the hazard for any other individual j:

$$\frac{h_i(t)}{h_j(t)} = \exp\{\beta'(x_i - x_j)\}.$$

This is an important result implying that the baseline hazard $h_0(t)$ can be omitted from estimation of β by using a partial likelihood function, which can be maximised. The likelihood function is 'partial', because it considers only cases with an observed event, and because it conditions out the time element. To construct a partial likelihood function it is necessary to have only the ranks of the event times, not their numerical values. So for the set of individuals at risk, $R(t_1)$, with ordered failure times, $t_1 < t_2 < ... < t_k$, the partial likelihood is given by

$$\boldsymbol{PL} = \prod_{i}^{k} \frac{\exp(\boldsymbol{\beta}'\boldsymbol{x}_i)}{\sum_{j \in \boldsymbol{R}(t_i)} \exp(\boldsymbol{\beta}'\boldsymbol{x}_j).}$$

The difficulty with this expression arises when the data contain tied failure times, i.e. failures that occur in one and the same period. This is the usual case in credit scoring, where account status is recorded on a monthly basis, so normally there is no information for rank ordering of the defaults that happened within one month.

There are several approaches to handle this problem. In the first, the 'exact' approach, Kalbfleisch and Prentice (1980) suggest that since the exact ordering of events is not known, all the possibilities should be included into the estimation, or *n!* orderings for *n* tied events. Whilst this approach provides the most accurate estimation, the computation becomes complicated with a large number of ties.

Two approximations were proposed. The Breslow approximation (Breslow (1974)) assumes that the ties occurred simultaneously and replaces *n* different orderings in the exact method by *n*! times the same ordering. The Efron approximation (Efron (1974)) assumes failures are sequential, but the orderings are not considered. Instead, a correction is made so that the first individual fails out of the full risk set, but each subsequent failure occurs from the risk set which excludes the average of all preceding failures plus this particular failure. Allison (1995) believes that the Efron approximation gives estimates closer to 'exact' values than the Breslow approximation.

Cox proposed to replace a continuous proportional hazards model with a discrete logistic model (Cox (1972)). It can be an appropriate approach for credit scoring when the monthly payments are to be made on the same date. This leads to a linear log-odds model:

$$\log(\frac{\boldsymbol{P}_{it}}{1-\boldsymbol{P}_{it}}) = \boldsymbol{\alpha}_t + \boldsymbol{\beta}'\boldsymbol{x},$$

where P_{it} is the conditional probability that individual *i* has event at time *t*, provided that the individual has not yet experienced the event and α_t is a set of constants that can vary between time periods.

To estimate this model, it is necessary to consider the probability that for any period of time with n events and k numbers at risk, events happened to these particular n individuals out of a possible set of k. It means one has to incorporate $k!/n!(k-n)!$ possibilities.

The following table gives estimation time in seconds that four different methods require for increasing sample sizes:

Table 5.1 Estimation time (sec) for different methods of handling the tied event times

	Sample size						
	100	200	400	600	800	1000	1200
Breslow	3	3	3	3	4	5	5
Efron	3	3	3	3	4	5	5
Exact	4	6	18	38	70	129	204
Discrete	3	4	6	9	12	19	26

Source: Allison (1995)

Stepanova (2001) found that there was almost no difference in the parameter estimates and no difference in the predictive ability between the discrete method and the Breslow approximation, although the log-likelihood statistic indicated a better fit for the discrete model. Following this result, the Breslow approximation is used in the current analysis.

Whilst the parametric AFT models provide more efficient estimation, PH has an advantage of robustness and flexibility. There are two important properties that follow from an arbitrary nature of the baseline hazard $h_0(t)$. First, $h_0(t)$ can be allowed to vary between different groups in the population. Second, it is possible to allow the covariates x to be time-dependent. The strength of association of a covariate with the default may vary with time, and the PH model can be extended to model time-dependency (Cox (1972); Stablein et al. (1981)).

5.3. Applications in credit scoring

The first attempt to apply survival analysis ideas to credit scoring was undertaken by Narain (1992). It was shown that estimates of the lifetime of a loan obtained from the exponential model can significantly improve the credit-granting decisions. Banasik et al. (1999) explored the issue further. Three models (exponential, Weibull and unspecified baseline proportional hazard) were compared to a standard logistic regression approach. The models were fitted to the fixed-term personal loan data and their performance was compared using two measures:

1) the ability to predict the probability of default in the first 12 months of the loan term;
2) the ability to predict the probability of default in the subsequent 12 months for the loans that survived through the first year.

For the first year the best performance was demonstrated by a non-parametric Cox model, followed closely by logistic regression and parametric proportional hazard models. For the second year the survival analysis methods performed worse than logistic regression, but this could be due to the fact that there were too few defaulters remaining in the risk set. The study applied the idea of competing risks in the credit scoring context, the competing risk was defined as early repayment of a loan. The proportional hazards models outperformed the logistic regression in predicting early repayments for the first year, but for the second year, logistic regression was the leader. It was suggested that the poor performance of the PH models in the second year was due to the fact that the ordering of risk of the event did not change with time. It was suggested that incorporating time-dependency might improve the results.

Stepanova and Thomas (2001); Stepanova and Thomas (2002); Stepanova (2001) give the most comprehensive investigation of all aspects associated with the application of the PH model in credit scoring. A new coarse-classing technique was proposed that was based on the PH approach and did not use any arbitrary time horizon. Methods of model diagnostics and tests for time-dependency were investigated and discussed in detail.

It was demonstrated that incorporating time-by-characteristic interactions improves the predictive ability of the model. Competing risks were incorporated into analysis, and methods were proposed to improve the prediction of early repayment. Finally, PH behavioural scoring techniques were developed that allowed the analyst to estimate the expected profit from the loans. It was shown that survival analysis is competitive with logistic regression in terms of predictive ability, and in addition, it offers advantages of providing estimates for profit scoring. Specifically, in terms of coarse-classification, the standard log-odds approach outperformed the PH method when the definition of bad was chosen to be 'early repayment before the end of a loan', i.e. the statistics measuring the separation between classes was based on the whole observed period. However, when the definition was changed to 'early repayment in the first 12 months', the PH approach showed better results than the log-odds one.

The performance of models was measured under the same two criteria as in Banasik et al. (1999). LR outperformed PH, with the PH model loosing a lot of power in the second year. Segmentation by the Loan Term significantly improved the performance of PH model for the second year, but still the LR was slightly better. When predicting default, segmented PH performed best of all for the first year, but segmentation did not improve the LR model. Non-segmented models showed similar performance. For the second year, LR was slightly better than PH, both with and without segmentation.

Behavioural scores were built for each month, predicting default in the remaining time period and incorporating the changing transactional information. The performance was measured again using two definitions of bad, but this time definition 2 was chosen to be ' default before the end of the loan', definition 1 was still 'default in the first 12 months'. The results for both definitions were very similar. LRB (Logistic Regression Behavioural) model performed better during the first year, while PHAB (Proportional Hazards Behavioural) model performance improved with time and after 2 years it was better than LRB.

Hand and Kelly (2001) used the idea of behavioural models based on a survival analysis approach in application to new products, when there are no historical data to develop a model on. The time to default was assumed to have an exponential distribution, which implied a constant hazard rate. At each time period, as new information on the performance of the accounts became available, the model was recalculated. It was shown that the survival analysis approach was superior to the logistic regression method, i.e. the former accepted less bads consistently throughout the whole period of observation. This was attributed to that fact that the logistic model tends to underestimate the true proportion of bads, since it assumes that the accounts which have not experienced default by a certain time point, are good. This may not be true because a certain number of these accounts will go bad later. The survival approach, however, assumes that all applicants will default eventually. Although this assumption seems to be unrealistic, for the purposes of modelling fixed term loans, it is possible to treat those who are predicted to default after the end of the loan term, as 'never going to be bad'.

All the above described research was based on fixed term credit products. In contrast, Till (2001) looked at the possibility of modelling delinquency in relation to revolving credit. However, the use of survival analysis in this study was limited to modelling the duration of stay in each delinquency state. Survival analysis was found to give similar results to a stationary Markov chain in estimating the mean duration spent in each state. The study also provided an investigation of the distribution form underlying the transactions. It was shown that the number of transactions up to time t can be modelled by a Negative Binomial, and times between transactions follow a Weibull distribution.

Baesens (2003) examined the application of neural networks for survival analysis modelling in credit scoring. Time to default and time to early repayment was considered. It was found that in predicting early repayment the neural networks outperformed the PH model, whilst in predicting default the superiority of the proposed approach was less evident.

5.4. Survival analysis applied to national and generic models

5.4.1. Data description.

Exactly the same national and generic samples were used as in Chapter 4 to allow for comparison of survival analysis models with logistic regression. The application data from Chapter 4 were merged with behavioural characteristics, which included the delinquency status, outstanding balance and credit limit, recorded on a monthly basis. The observation period was 25 months from November 1998 until December 2000. Exactly the same allocation into training and hold-out datasets was kept.

Survival analysis predicts the time until a certain event. For the purpose of this analysis, the event was defined as 'first time 2 months in arrears'. The accounts that did not experience the event were considered censored. The life of the account was measured from the month it was opened until the event or until the account was closed. In the latter case, the account was considered to be censored. All other cases were censored at the end of observation period.

The data used in the analysis in this study contains right-censored observations. As Table 5.2 indicates, the majority of cases have right censoring because they did not get into the default category and kept the card account open up to and including November 2000, which was the last period when the account status was recorded. These cases, however, can be considered as randomly censored since although the observation ended at a fixed time point, the start of the process was not the same time point for all accounts, so the actual survival times can be considered random. There are some customers that have closed the account before the last recorded period. However, there are no reasons to suspect that this can be related to the likelihood of the default, so it can be assumed that it is random censoring.

Table 5.2. Different types of censoring by country.

	Non-censored	Randomly right-censored	Time right-censored
Belgium	3090	521	22679
The Netherlands	11213	5719	65305
Germany	8909	2982	63957

5.4.2. Survival patterns by country

Before modelling the relationship between application characteristics and survival times, some exploration of survival patterns was done for each country. This involved:

1) fitting KM estimates of the survival curves, as described in Section 5.2.4. (Figure 5.1);
2) examining the negative log-survival (LS) plots (Figure 5.2) and a plot of $log[-log\hat{S}(t)]$ versus $log\ t$ (LLS) (Figure 5.3);
3) examining the hazard rate plot, with estimates of the hazard function obtained as described in Section 5.2.4. Figure 5.4 presents the 95% confidence intervals for the hazard rate for each country.

The SDF plots (Figure 5.1) overlap for Belgium and Germany, with SDF for the Netherlands decreasing faster and the difference getting larger with time. The Log-rank test and Wilcoxon test showed that survival curves were significantly different for the three countries (Table 5.3).

Table 5.3. Test of equality of SDF between 3 countries

Test	**Chi-Square**	**Degrees of freedom**	**Pr > Chi-Square**
Log-Rank	131.2216	2	0.0001
Wilcoxon	97.8208	2	0.0001

To check whether the hazard is constant over time *–log (S)* is plotted against the time (Figure 5.2). The behaviour between 2-5 months is different between the countries. Germany follows nearly a straight line. For Belgium the chances of survival decrease at a faster rate until month 3 before levelling out. The Netherlands shows a smooth (but not constant) decline in survival probability. After month 5, the graphs seem to follow approximately a straight line, which implies a constant hazard (see Section 5.2.4) and suggests that exponential distribution can be a suitable fit, although a slight downward curvature can be observed. As in the previous graph, whilst for Belgium and Germany the graphs overlap, the Netherlands exhibits the lower chances of survival.

Figure 5.1 Survival Distribution Function by country

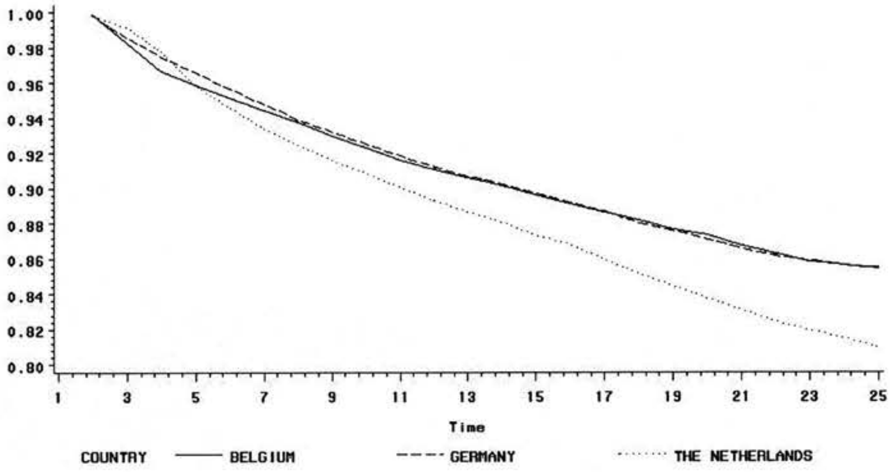


Figure 5.2 Negative log SDF against time by country

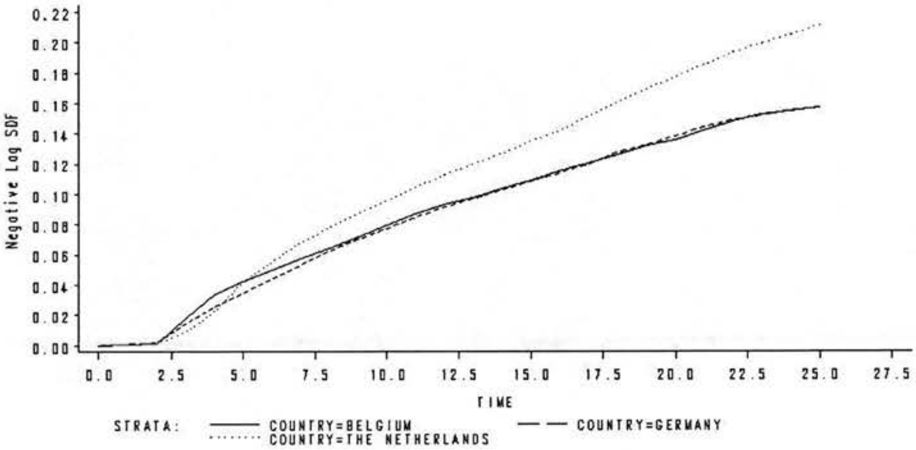
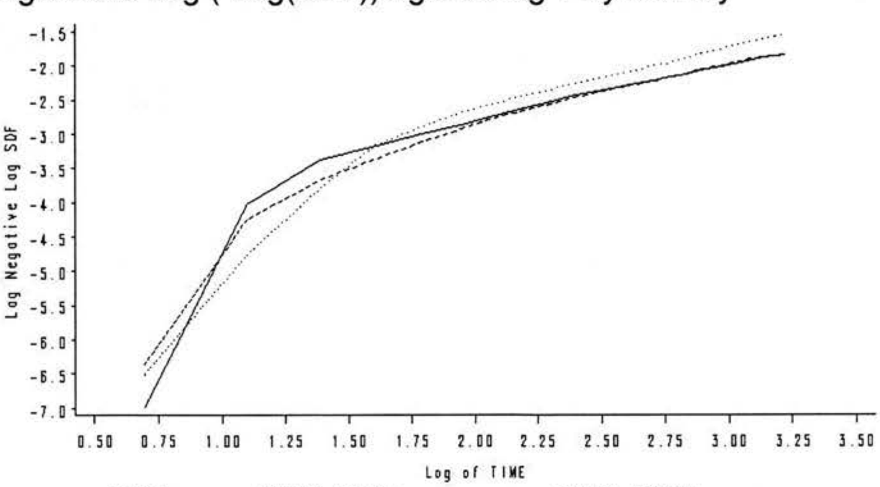


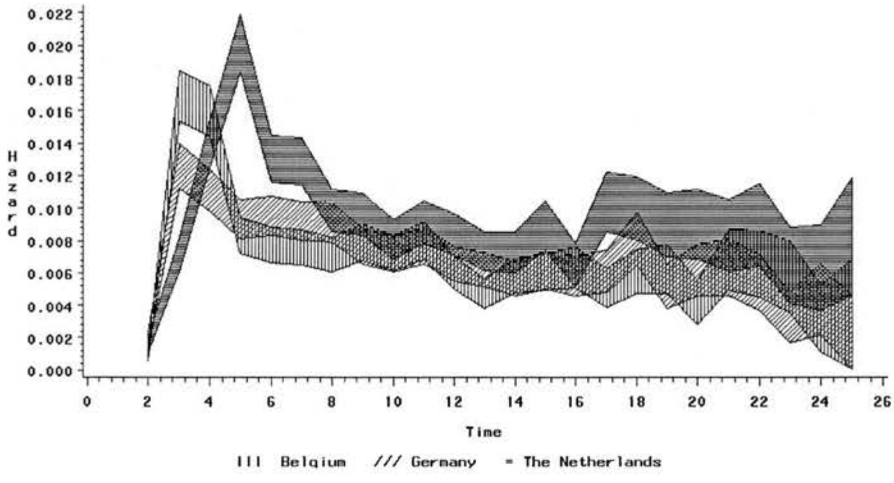
Figure 5.3 Log (-Log(SDF)) against log T by country



A plot of $log[-log\hat{S}(t)]$ versus $log\ t$ (LLS) shows whether the hazard rate is changing monotonically (Figure 5.3). A downward curve is observed for the Netherlands. It is possible to distinguish two parts with approximately straight lines for Germany, and three parts for Belgium. Overall, the investigation of the plot suggests that the hazard rate is not changing monotonically, but for Germany and Belgium it is possible to split the total observation period into parts with monotonic hazard rate.

The hazard plots (Figure 5.4) for all three countries increase rapidly in the first months of the account life and then decrease towards an asymptote. This supports a conventional wisdom - 'if they go bad, they go bad early' (Banasik et al. (1999)), and is in line with results for fixed term loans (see Stepanova (2001)). However, 'early' means 2 months for Belgium and Germany, and 5 months for the Netherlands. The height of peaks also differs between the countries, with Germany being the least risky in terms of early defaulters and the Netherlands being the most risky. After 9 months the confidence intervals overlap, but the Netherlands remains slightly higher and shows a slight increase at the end of the observation period. However, this increase could be due to the increased variability, as indicated by a wider confidence interval, because the risk set becomes small.

Figure 5.4 Hazard function with 95% confidence intervals



The difference in behaviour between accounts from the Netherlands and the other two nations may in part be due to a higher proportion of accounts with deferred payment schemes in the Netherlands. Table 5.4 gives the breakdown by types of contract for each country. For Belgium deferred options constitute only 2.16% of the accepted accounts. In Germany this percentage is 10.52%. Still it is possible to say that for Belgium and Germany the populations are fairly homogeneous in terms of contract terms, with nearly 90% falling into 'Budget type of agreement, customer contacted by phone' type. In contrast to them, the proportion of deferred payment schemes in the Netherlands is 41.76% and constitutes the largest group of contract terms.

Table 5.4 Percentage breakdown of different agreement types by country

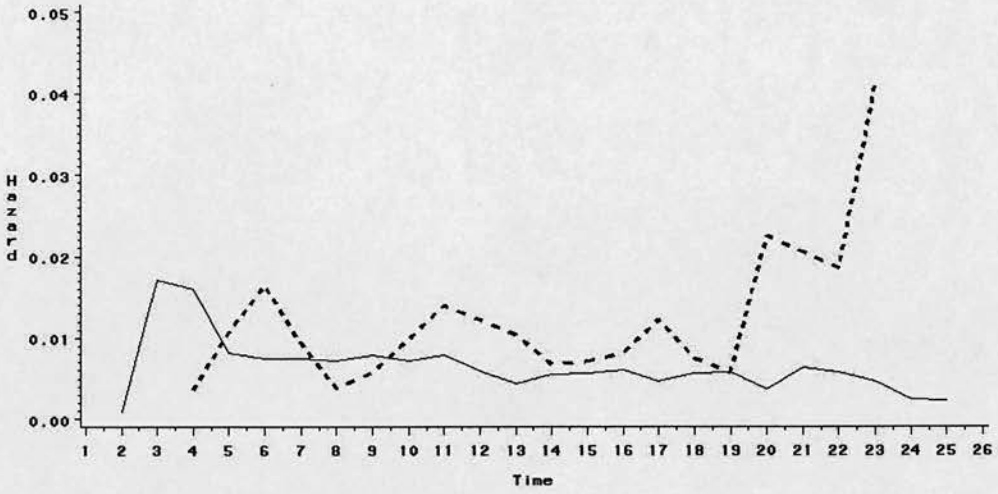
Agreement Type	Bads	Goods	Rejected	Total accepted	Total
Belgium					
Total Deferred	2.30%	2.14%	0.66%	2.16%	1.55%
Total Budget	88.32%	88.18%	96.48%	88.19%	91.56%
Total Other	9.39%	9.68%	2.87%	9.65%	6.90%
Germany					
Total Deferred	12.20%	10.30%	9.34%	10.52%	9.99%
Total Budget	85.57%	88.10%	89.31%	87.80%	88.48%
Total Other	2.23%	1.60%	1.35%	1.68%	1.53%
The Netherlands					
Total Deferred	48.61%	40.68%	36.16%	41.76%	39.67%
Total Budget	29.87%	29.17%	37.66%	29.27%	32.40%
Total Other	21.52%	30.14%	26.18%	28.97%	27.93%

Figure 5.5 gives hazard plots for deferred and non-deferred options by country. For the Netherlands deferred options are riskier than non-deferred and cause the peak at month 5. Non-deferred options have an almost constant hazard rate, and contribute towards the slight increase in the hazard rate in the end of the observation period.

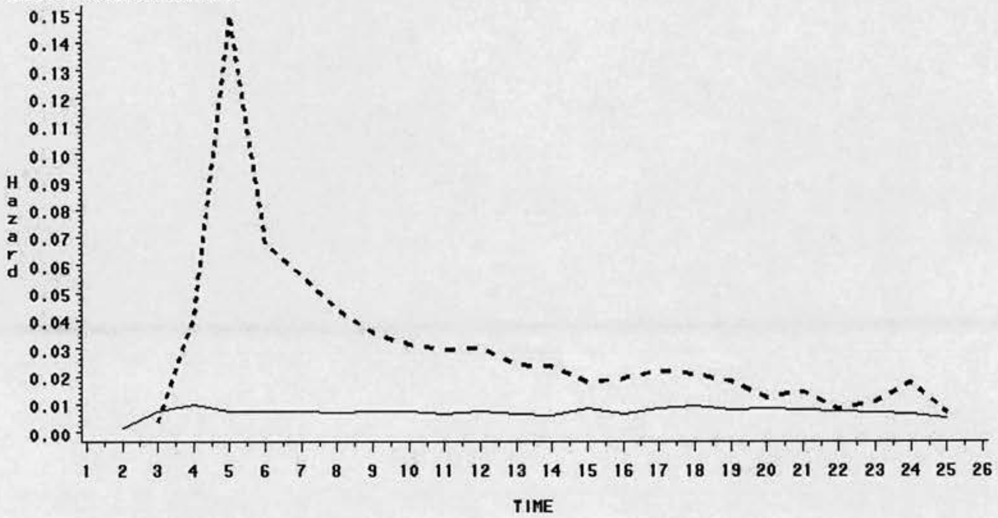
In Germany and Belgium the deferred options are also riskier than non-deferred although the difference is not as striking as in the Netherlands. Such a difference between the countries may be explained by the number of deferred accounts. In the Netherlands the number of these accounts is large enough to accommodate probably, the majority of applicants with financial problems, who would obviously prefer to pay later. In Germany and Belgium such applicants have to choose some other options, and hence they are more evenly split between different agreement types.

Figure 5.5 Deferred/Non-deferred agreements by country

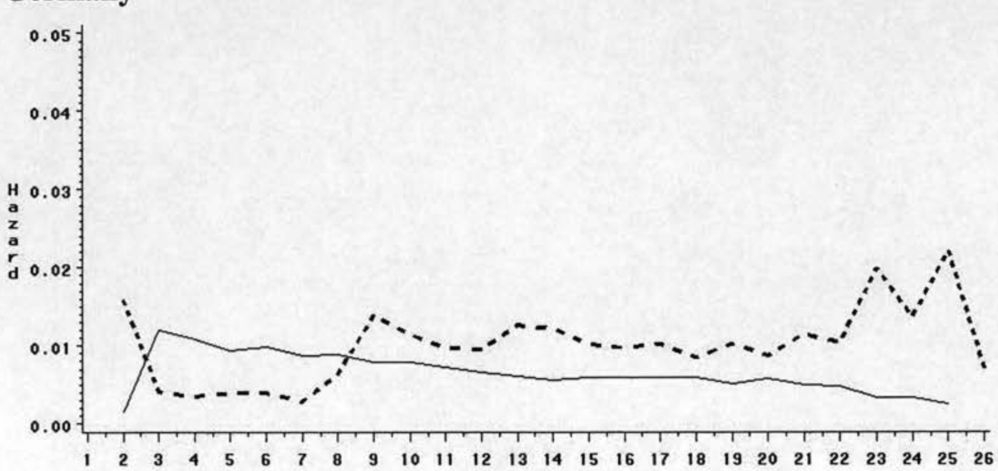
Belgium



The Netherlands

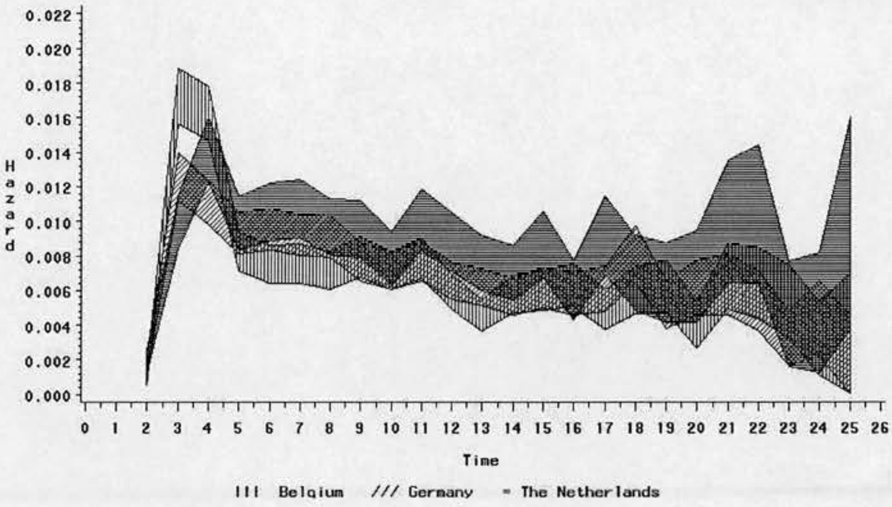


Germany



The hazard plot with deferred options removed is given in Figure 5.6. The patterns for Belgium and Germany remain unchanged, which is to be expected, since the proportion of deferred schemes is not large in these countries. For the Netherlands, it is obvious that the relatively large peak at month 5 is due to deferred accounts.

Figure 5.6 Hazard function with 95% confidence intervals (non-deferred payment schemes)



The analysis proceeded with accounts of all agreement types included for the Netherlands, because it is a significant proportion of data and removing it will eliminate a lot of information. However, the agreement type was not included into the set of predictor variables, since there was a large number of unique types for the Netherlands. Including these types as a covariate would be equivalent to including a country indicator variable, so the model would be not exactly 'generic'. This creates an unobserved heterogeneity situation, but the conditions remain equal for all different methods of survival analysis and LR used in the analysis, so the comparison is still valid.

Tests of equality over countries with deferred schemes removed (Table 5.5) still showed significant values, although the chi-square statistics became less in magnitude.

Table 5.5 Test of equality between countries (non-deferred payment schemes)

Test	Chi-Square	Degrees of freedom	Pr > Chi-Square
Log-Rank	27.1494	2	0.0001
Wilcoxon	17.5926	2	0.0002

Overall, Figures 5.1-5.6 suggest that apart from the first 3 months, the exponential distribution can be a suitable approximation, while with the first 3 months log-normal or log-logistic seems to be more appropriate.

5.4.3. Coarse-classification using the log-odds and survival analysis techniques

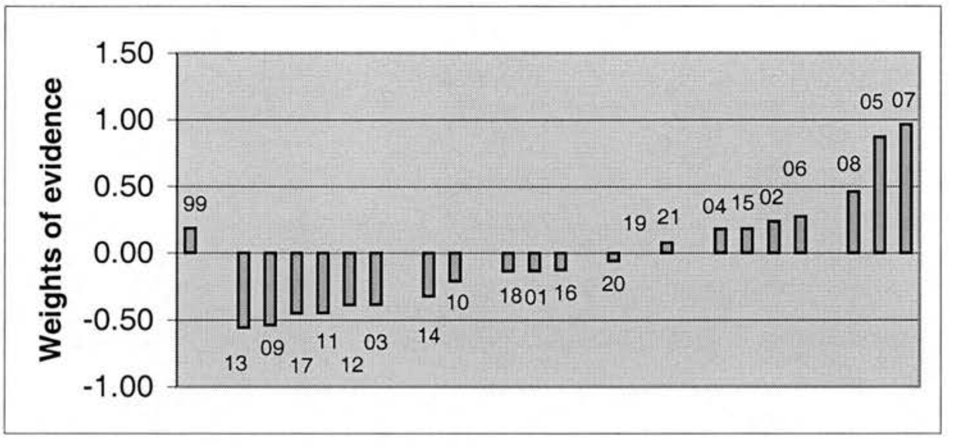
Before fitting regression models to the data, the attributes were grouped together to form binary variables like in Chapter 4. But the rationale for grouping the attributes was different and was based on the approach proposed by Stepanova (2001):

1. The continuous characteristics were fine-classed into 20 groups, each group containing 5% of data. Attributes of discrete characteristics were fine-classed on the basis of their meaning. For example, for the 'Goods code', attributes 'Bedroom furniture' and 'Living-room furniture' were combined into 'Furniture' group. For discrete characteristics with a small number of attributes, no fine-classing was made.
2. Then proportional hazards model was fitted to fine classes, and the parameter estimates were plotted.
3. Fine classes with similar parameter values were combined into coarse classes.

Figures 5.7 and 5.8. show examples of coarse-classification based on proportional hazards and on the log-odds approach from Chapter 4. Both approaches suggest identical groupings. This is different from the results reported by Stepanova (2001). This can be attributed to the fact that here the data relates to revolving credit as opposed to fixed term personal loans and there is no pronounced time structure, as was shown in exploratory analysis, apart from early peaks in defaults.

Since there was no difference between the log-odds and survival analysis coarse-classes, the binary variables obtained in Chapter 4 were used in the subsequent analysis. So in cases when there is no pronounced time dependency the coarse-classification can be done using weights of evidence. It may be an advantage with some computer packages that do not convert categorical variables into binary ones automatically.

Fig. 5.7 Example of coarse-classification of 'Business Type'



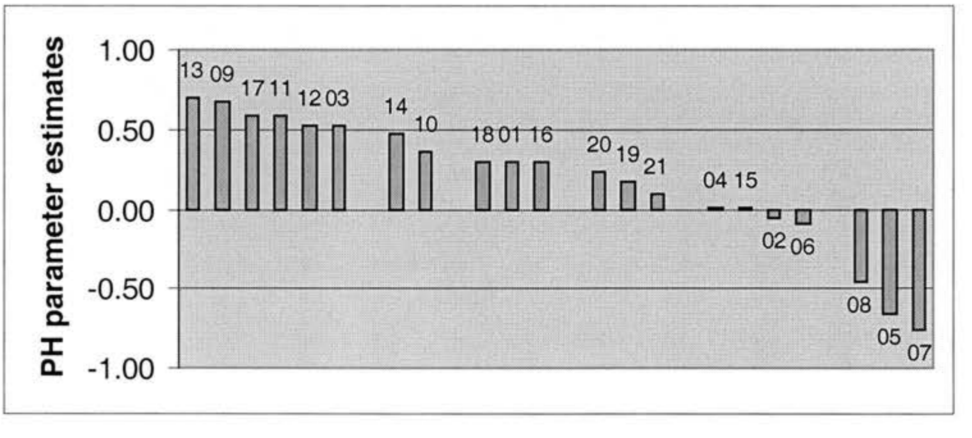
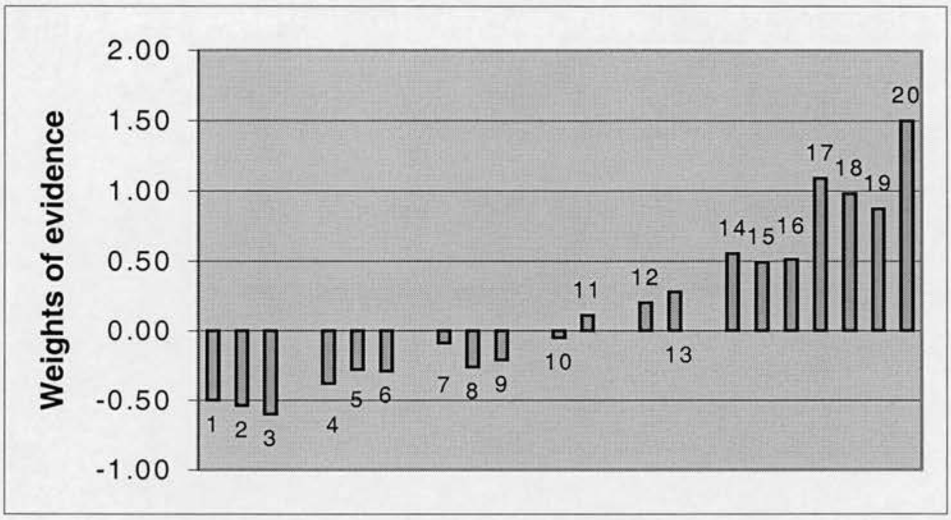
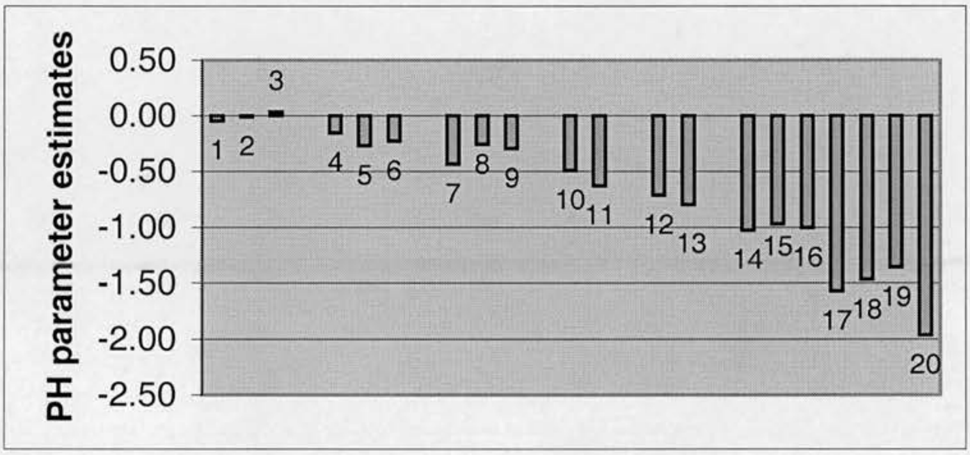


Fig. 5.8 Example of coarse-classification of 'Time at address'





5.4.4. Survival analysis compared to logistic regression

The preliminary exploration of data (hazard plots) suggested a number of possible approaches: log-normal and log-logistic distributions seemed to be most suitable for modelling the delinquency data, although one should not discard the Weibull family distributions. At the same time the country plots looked roughly proportional, suggesting that PH models would be appropriate.

So exponential, Weibull, log-normal, log-logistic distributions for parametric models were tried and Cox proportional hazards model with unspecified baseline hazard function. The model fit was measured by -LogLikelihood, with lower magnitude values indicating the better fit. This measure indicates how well the model describes the data but in relative, not absolute terms. It is used to compare nested models. Models are nested if one is a special case of another, in our case the exponential, Weibull, log-normal models are nested within the generalised gamma. This allows one to make tests of significance for the difference in model fit. Log-logistic, Logistic and Cox proportional hazards model cannot be compared directly on the basis of this measure, since they are different, not nested models. However, the judgements can be made based on the predictive ability of models, which was measured by the area under the ROC curve and error rate.

The predictive ability of the models was measured on hold-out samples. The score produced by the model provided estimates of log T. From this the estimates of the 'survival' time were obtained or estimates of time that the customer is going to use the card without going into default. This time was rank-ordered, the shorter times were considered to be 'bad'. The cut-off was chosen to match the number of actual bads in the hold-out sample, as in Chapter 4. An alternative way would be to obtain the probability of 'surviving' until some specified period, which is equivalent to the way predictions are generated from logistic regression. However, the advantage of survival analysis consists in its ability to produce predictions for several time periods from the same model, which logistic regression cannot do. For PH models the estimate of the hazard was used as a score.

The results from different survival analysis models and logistic regression are reported in Table 5.6. AFT models should be compared with proportional hazards and logistic full models (not stepwise), since there is no variable selection procedures incorporated in SAS for AFT models.

The differences in model fit were highly significant (Table 5.7). In terms of model fit, the leader was the gamma distribution followed by the log-normal for all three countries. However, this did not translate into superior prediction results. In fact, gamma gives the worst prediction, which proves a well-known fact that better fit does not necessarily mean better prediction (Hosmer and Lemeshow (2000)).

Table 5. 6. Survival analysis and logistic regression models by country

Model	Log L	AUROC	Error rate
Belgium			
Exponential	-7842	0.7121	16.74%
Weibull	-7811	0.7119	16.78%
Loglogistic	-7780	0.7123	16.74%
Lognormal	-7708	0.7122	16.90%
Gamma	-7690	0.7116	16.90%
PH	-20093	0.7122	16.72%
PH stepwise	-20480	0.7071	16.74%
Logistic	-5935	0.7129	16.92%
Logistic stepwise	-5990	0.7074	16.96%
The Netherlands			
Exponential	-24656	0.7802	16.50%
Weibull	-24185	0.7797	16.50%
Loglogistic	-23870	0.7800	16.50%
Lognormal	-23700	0.7801	16.54%
Gamma	-23693	0.7801	16.54%
PH	-79429	0.7802	16.50%
PH stepwise	-79435	0.7798	16.56%
Logistic	-19229	0.7814	16.50%
Logistic stepwise	-19216	0.7804	16.50%
Germany			
Exponential	-22531	0.7408	15.78%
Weibull	-22380	0.7405	15.78%
Loglogistic	-22269	0.7406	15.78%
Lognormal	-22062	0.7406	15.86%
Gamma	-22016	0.7404	15.88%
PH	-65386	0.7412	15.76%
PH stepwise	-38588	0.7395	15.78%
Logistic	-17389	0.7417	15.74%
Logistic stepwise	-17421	0.7394	15.74%

Table 5.7 Model fit test of significance based on Log Likelihood

	Chi-sq	df	p>0
Belgium			
Exponential vs Weibull	44	1	2.76018E-11
Weibull vs Gamma	429	1	2.4511E-95
Lognormal vs Gamma	143	1	4.61618E-33
Exponential vs Gamma	473	2	1.5012E-103
The Netherlands			
Exponential vs Weibull	941	1	1.4026E-206
Webull vs Gamma	985	1	3.5089E-216
Lognormal vs Gamma	13	1	0.000280096
Exponential vs Gamma	1926	2	0
Germany			
Exponential vs Weibull	49	1	3.00302E-12
Webull vs Gamma	1177	1	6.7201E-258
Lognormal vs Gamma	335	1	7.43031E-75
Exponential vs Gamma	1225	2	7.7443E-267

Nevertheless, all distributions are amazingly close in predictive accuracy. So given the preference for a more parsimonious and therefore robust model, the exponential distribution would be most suitable from parametric models. At the same time proportional hazards and logistic regression models give identical results. Since there is no or little difference in predictive accuracy, the decision as to which approach to use should be based on additional properties that a certain method can provide.

One can argue that there may be periods within 25 months when certain models give superior prediction. This argument would apply to AFT survival models since they allow for modelling the changing hazard rates between different groups over time, and therefore will produce different ranking of borrowers over time. To test this proposition the survival models were applied to two alternative definitions of bad. First, those that defaulted within the first six months were considered to be bad, and the rest were treated as good. Second, those that defaulted within first twelve months were classified as bad, the remaining customers in the hold-out sample were considered good.

The results are given in Table 5.8. It is possible to say that log-logistic, log-normal and gamma models give slightly superior prediction for Belgium and the Netherlands and especially for 'Default in 6 months'. This is in line with hazard plots (Figure 5.4) when the peaks in hazards during the first 6 months were observed, so if one would expect some difference in results that would be during the first months. The hazard peak for Germany was least pronounced, and hence there is no marked difference between the predictive ability of survival models. It should be noted though that even for Belgium and the Netherlands the differences are marginal and do not give enough grounds to conclude that log-logistic, log-normal or gamma should be preferred to more robust exponential and PH models. Perhaps, the superiority of AFT models may be more visible in different credit scoring applications (e.g. insurance) where there is a more pronounced time structure.

Table 5.8 Survival models tested for alternative definitions of default

Model	**Default in 6 months**		**Default in 12 months**	
	AUROC	**Error rate**	**AUROC**	**Error rate**
Belgium				
Exponential	0.7225	8.30%	0.7225	13.20%
Weibull	0.7224	8.26%	0.7225	13.26%
Loglogistic	0.7232	8.30%	0.7229	13.26%
Lognormal	0.7235	8.22%	0.7223	13.40%
Gamma	0.7234	8.12%	0.7218	13.44%
PH	0.7226	8.32%	0.7224	13.26%
The Netherlands				
Exponential	0.8241	7.36%	0.8036	12.24%
Weibull	0.8249	7.38%	0.8037	12.20%
Loglogistic	0.8255	7.24%	0.8046	12.22%
Lognormal	0.8262	7.26%	0.8052	12.24%
Gamma	0.8263	7.26%	0.8053	12.26%
PH	0.8242	7.26%	0.8036	12.22%
Germany				
Exponential	0.7649	7.02%	0.7483	12.56%
Weibull	0.7639	7.02%	0.7482	12.54%
Loglogistic	0.7639	7.00%	0.7483	12.56%
Lognormal	0.7637	7.00%	0.7482	12.52%
Gamma	0.7635	6.98%	0.7480	12.50%
PH	0.7642	7.04%	0.7484	12.58%

5.4.5 Performance of generic models

Three basic types of models were selected for comparison between different modelling approaches in building generic models: the exponential model to represent the family of parametric models, the PH non-parametric model and logistic regression. In this section only full models (containing all common variables with all categories, both significant and insignificant) are considered, so that comparison could be made with the exponential AFT model.

The generic models were applied to four hold-out samples: a generic one that consists of residents of the three countries, and three national samples, the predictive performance was measured by AUROC (Table 5.9a) and error rate (Table 5.9b). The national LR models are used as a benchmark for comparison of performance of generic models on the corresponding national hold-out sample. The best predicting generic model is marked in bold.

Table 5.9a) Predictive performance of generic models. AUROC

Model	**Hold-out sample**			
	Belgium	**The Netherlands**	**Germany**	**Generic**
National LR	0.7129	0.7814	0.7417	N/A
Generic models, no country indicators				
Generic LR	0.7027	0.7790	**0.7316**	0.7465
Generic EXP	0.7029	0.7777	0.7315	0.7464
Generic PH	0.7031	0.7777	0.7315	0.7463
Generic models, country indicators				
Generic LR	0.7047	**0.7799**	**0.7316**	**0.7477**
Generic EXP	0.7047	0.7785	0.7315	0.7475
Generic PH	**0.7048**	0.7785	0.7315	0.7475
Generic PH models, stratification / baseline				
Stratified PH, no baseline	**0.7048**	0.7786	0.7314	0.7438
Stratified PH, baseline	**0.7048**	0.7786	0.7314	0.7469
Generic PH, baseline	0.7031	0.7777	0.7315	0.7473

Table 5.9b) Predictive performance of generic models. Error rate

Model	Hold-out sample			
	Belgium	**The Netherlands**	**Germany**	**Generic**
National LR	16.92%	16.50%	15.74%	N/A
Generic models, no country indicators				
Generic LR	17.02%	16.68%	**16.16%**	**16.48%**
Generic EXP	17.06%	16.68%	**16.16%**	16.54%
Generic PH	17.08%	16.68%	**16.16%**	16.54%
Generic models, country indicators				
Generic LR	16.90%	**16.60%**	**16.16%**	16.58%
Generic EXP	16.80%	16.62%	16.18%	16.62%
Generic PH	**16.72%**	16.62%	16.18%	16.64%
Generic PH models, stratification / country baseline				
Stratified PH, no baseline	16.74%	16.62%	16.18%	16.50%
Stratified PH, baseline	16.74%	16.62%	16.18%	16.64%
Generic PH, baseline	17.08%	16.68%	**16.16%**	16.66%

The customized national models still give slightly better prediction, while three generic models (LR,EXP, PH) demonstrate an amazingly close performance. One can argue that logistic regression is slightly superior, but again the difference is marginal. The inclusion of the country indicator leads to a very modest improvement in predictive ability of models, if any at all. The error rate for the generic hold-out sample even demonstrates a slight loss in the discriminating power.

An important property of the PH model is its ability to take into account different subpopulations that may exist within the data, while producing a single set of parameter estimates that does not include a subpopulation indicator. This process is called stratification and is commonly used for subpopulations that violate the proportionality assumption. It should be noted, though, that in order to use this property one needs to know the applicant's country of residence. Still it will be useful in situations when 'nationality' is not legally forbidden, and lenders can replace several national models with a generic one that accounts for different subpopulations due to stratification.

The following model is fitted to data:

$$\log \boldsymbol{h(t)}=\alpha_z(\boldsymbol{t}) + \boldsymbol{\beta'x} ,$$

where z is a subpopulation indicator. In this way the hazard function is allowed to vary between specified groups. The separate partial likelihood functions are constructed for each level of z. The functions are multiplied together, and then the values of β are estimated that maximise the combined PL function. This method presents a half way option between generic and customised models: to estimate the parameters the country indicator is required, but the subsequent scoring of new accounts can be done without making a distinction between the countries.

The stratified PH model has been fitted to the aggregated generic dataset (see Table 5.9). AUROC shows no improvement on the generic hold-out sample, but when tested separately on the national samples, there is some increase in AUROC for Belgium and the Netherlands, although not a dramatic one. Error rate demonstrates some superiority of the stratified approach if compared to other models, including the logistic regression. For Belgium the stratified model shows the error rate even lower than the national LR model.

Further, although a baseline hazard function cancels out of the PH partial likelihood model for β, as was shown in Section 5.2.5.2, a baseline survivor function can be obtained by non-parametric methods taking into account the estimates of β from a fitted PH model, see Kalbfleisch and Prentice (1980).

The PH model can be written as:

$$S(t)=[S_0(t)]^{\exp(\beta'x)}$$

where $S_0(t)$ is a baseline survivor function. With the baseline survivor function estimate it is possible to generate probabilities of surviving to time T, and to use different country baseline hazard functions when scoring the new applications.

When the country baseline estimates were raised to the power of the generic stratified and non-stratified PH scores, this showed some improvement in AUROC, but on the other hand, deterioration in error rate. The national measures remained unchanged since this monotonic transformation does not change the ranking of applications.

In general, one can conclude that stratification brings some benefit but not a convincing one. Such modest improvements can be attributed to the fact that the hazards between the countries are roughly proportionate, so there is little scope for stratification to enhance the predictive performance. Overall, generic models, as given in Table 5.9, do not demonstrate a notable difference in predictive accuracy, which is further confirmed by inconsistencies between two measures: AUROC and error rate.

5.4.6 Predicting early defaulters

The examination of the hazard functions in Figure 5.4 shows that the hazard is the highest at the early stages of the credit life. It would be of interest to investigate whether any of the generic models developed in the previous section are superior in predicting 'early' defaulters, the group of applicants which is, probably, least wanted by credit grantors.

We defined an 'early' period as the time period of the first 8 months because:

1) after month 8 the hazard curves remain roughly constant;
2) nearly 50% of bad credits that were observed during 25 months, defaulted within the first 8 months.

So the cardholders that missed two consecutive payments during 8 months were classified as 'bad', and survival analysis and logistic regression generic models were tested for their ability to predict this category of applicants. For the survival analysis methods a change of the definition of 'bad' did not present a problem, since both AFT and PH models developed in the previous section could be used to predict the probability of survival to any specific time period. The logistic regression model could only generate a probability of defaulting in 25 months, so a different LR model was estimated that modelled the probability of defaulting in 8 months. However, the '25 months' models were also included in the experiment, since one can argue that applicants with higher probability of default in 25 months will default earlier than those with lower chances of default in 25 months.

The results are given in Table 5.10.

Table 5.10a) Predicting 'early' defaulters. AUROC

Model	Hold-out sample			
	Belgium	The Netherlands	Germany	Generic
Generic models, no country indicators, Section 5.4.6				
LR (25m)	0.7050	0.8105	0.7499	0.7620
EXP	0.7062	0.8096	**0.7503**	**0.7625**
PH	**0.7064**	0.8095	**0.7503**	0.7623
Generic models, re-estimated				
LR (8m)	0.7009	**0.8120**	0.7493	0.7603
Time-dependent PH (25m)	0.7054	0.8101	**0.7503**	0.7621
Time-dependent PH (8m)	0.7016	0.8101	0.7491	0.7598

Table 5.10b) Predicting 'early' defaulters. Error rate

Model	Hold-out sample			
	Belgium	The Netherlands	Germany	Generic
Generic models, no country indicators, Section 5.4.6				
LR (25m)	10.00%	9.36%	9.34%	9.66%
EXP	10.06%	9.38%	9.34%	9.66%
PH	10.06%	9.38%	9.34%	9.66%
Generic models, re-estimated				
LR (8m)	9.98%	**9.18%**	**9.26%**	**9.48%**
Time-dependent PH (25m)	10.02%	9.26%	9.36%	9.60%
Time-dependent PH (8m)	**9.92%**	9.22%	9.30%	9.50%

The PH model can be extended to model time-dependency (Cox (1972), Stablein et al. (1981)), and this is another common way to handle the variables that show the non-proportionate behaviour. If the effect of the variables changes with time, this can be included into the model as a variable-by-time interaction:

$\log \boldsymbol{h}\,(t)=\alpha_z(t) + \beta'x + \gamma'xt$, or $\log \boldsymbol{h}\,(t)=\alpha_z(t) + (\beta' + \gamma' t)x$.

So if γ is positive, the effect of a variable increases linearly with time, if it is negative, the effect decreases with time. First, t was allowed to take values from 1 to 25 months, this is marked in Table 5.10 as PH (25m). With this specification the estimate of γ shows how much the effect increases/decreases every month. The parameter estimates are given in Appendix, A24.

To evaluate the chances of the borrowers to be bad in the first 8 months, the following score was used:

$(\beta' + \gamma' t)x$, where $t = 8$.

Second, the time variable was dichotomised, with $t = 1$ representing months from 1 to 8, this is marked as PH (8m) in Table 5.10. As can be seen from Table 5.10 incorporating time-dependency has a limited effect on prediction. This can be explained by weak time effects. In fact, as the output in Appendix A24 shows, there are only 6 variable-by-time interactions that are significant at 5% level. This also explains why the exponential model with its 'lack of memory' predicts well on this dataset.

Overall, there is no evidence that any of the models tested consistently outperformed the others.

5.4 Conclusions

This Chapter presented the first cross-country comparison of survival analysis applied to modelling time to default. Besides, the analysis was conducted on a revolving credit product, which is different from previous studies that analysed fixed term loans. In addition, different approaches within survival analysis were investigated.

The analysis of this Chapter supports the previous findings that survival analysis is competitive with the logistic regression in application to credit scoring. The comparison of several approaches within survival analysis showed that there was little difference in classification accuracy between the parametric AFT, non-parametric Cox proportional hazards models and logistic regression.

So from the point of view of prediction, any of the tested approaches can be used. However, survival analysis offers a number of benefits that potentially makes it superior if compared to logistic regression. First, predictions can be generated that give the probability of 'survival' in specific time period, without the necessity of re-estimating the model to fit this time period, as is the case with logistic regression.

Second, by means of stratification it is possible to account for differences between subgroups in the population without including the subgroup indicator explicitly into the model. And third, it is possible to incorporate the variable effects that change with time.

It has been shown in this Chapter that countries used in analysis demonstrate different survival patterns, with Belgium and Germany being close to each other, whilst the Netherlands exhibiting the higher chances to default. The examination of the negative log-survival curves (Figure 5.2) suggests that the exponential model is a suitable approximation for all three countries.

But in spite of the observed differences, generic models demonstrate only a very slight loss in predictive power as compared to national models, in line with results of Chapter 4. Again there is practically no difference between generic logistic regression, exponential AFT and semi-parametric Cox proportional hazards models. Because the hazard functions are roughly proportionate between the countries, stratification has a very limited effect. So does the inclusion of variable-by-time interactions, which can be explained by weak time effects.

Weak time effects together with the fact that the exponential model with its 'lack of memory' property fits the data well suggest that revolving credit may be different from the fixed term loans in that it may have a more random character. However, this proposition needs to be tested on other datasets.

It has been shown that survival analysis is suitable for building generic models for predicting default. The next Chapter will investigate the possibility of using survival analysis to predict time to the next purchase in generic models.

Chapter 6. Predicting time to the second purchase

6.1 Introduction

In the previous Chapter it was shown that national and generic survival analysis models are competitive with the logistic regression models in predicting default. Whilst predictive accuracy of these models is very similar, survival analysis offers an additional advantage of having a time perspective, thus laying the foundation for profit scoring.

However, in the context of revolving credit it is also important to consider the usage of the credit product - a retail card in this case- when trying to estimate profit. For example, there may be a good customer that never goes into arrears, but at the same time makes only one purchase, repays it quickly and never uses the card again. On the contrary, a bad customer (according to the traditional definition) may fall behind with payments, but then may repay everything subsequently and make several new purchases on the card. The first customer, though presenting a lower risk of default, would also yield less profit than the second customer.

The objective of this Chapter is to explore the possibility of predicting the future usage of the card. Usage can be described in a number of ways: one may consider a binary outcome, whether a customer uses the card after the first purchase again or not; or one may try to predict the number or total value of purchases.

With fierce competition in both the lending and retailing sectors, it is the time component that becomes of extreme importance. Knowledge of projected time to the next purchase can be helpful in deciding how quickly the lender's intervention is required in order to retain a potentially profitable customer. That is why this Chapter considers the time to the second purchase, in order to provide insight into an appropriate time scale for the actions that have to be taken by the lender. This investigation is complemented by an exploration of national differences in purchasing behaviour and the possibility of accommodating these differences by a single generic model.

The analysis presented in this Chapter investigates the incremental roles of information which becomes available at various stages of the credit management process in explaining the time taken until the borrower makes a second purchase. It is established that the best prediction can be achieved by using a combination of application and behavioural variables and that characteristics of the first purchase and the remaining credit available are the most powerful predictors. The significance of behavioural data increases with time, whereas the application characteristics gradually loose some of their discriminating power. So behavioural scoring considerably enhances the ability to predict when a second purchase will be made. Lenders could use this kind of 'purchase scoring' to assess the chances of the purchase in the next month, and therefore forecast their cash flow and adjust customers' credit limits.

The Chapter is structured in the following way: Section 2 summarises previous research on the prediction of usage, Section 3 describes the approach taken and the data available for analysis. Section 4 presents the results on the country basis and investigates the value of different levels of information in explaining time to the next purchase: personal characteristics, characteristics of the first product purchased and transactional behavioural data. Section 5 reports on the predictive performance of the national models. Section 6 investigates the differences in the purchasing behaviour of the customers who go into delinquency and those that do not. Section 7 compares the predictive performance of the generic models to the national alternatives. Finally, Section 8 concludes.

6.2. Literature review on predicting usage

Whilst predicting default is a well-established technique in credit scoring, estimation of usage is far less common. However, estimated usage can be extremely important in selecting an appropriate credit limit and this will have a significant effect on the customer's spending behaviour. Normally the credit limit is initially assigned on the basis of the applicant's risk score and income (Thomas et al. (2002)). The credit limit is then reviewed in one year's time on the basis of a behavioural score, which is as a rule risk oriented (Oxley (2003)).

The number of studies that have investigated the propensity to use a credit product is relatively limited. Research has concentrated predominantly on discrimination between users and non-users of credit cards (Lindley et al. (1989), Crook et al. (1992), White (1975), Carow and Staten (1999)) or on predicting different levels of usage (Volker (1982), Hirschman (1982), Banasik et al. (2001)).

It was found that the following determinants of the customer's propensity to repeatedly use a revolving credit product were most important:

- age (Crook et al. (1992), White (1975), Carow and Staten (1999), Volker (1982)),
- race (Lindley et al. (1989), White (1975)),
- marital status (White (1975), Banasik et al. (2001)),
- gender (White (1975)),
- education (Carow and Staten (1999)),
- skilled/unskilled indicator (Volker (1982)),
- customer's views and attitude to different payment methods (Hirschman (1982)),
- residential status (Crook et al. (1992), Banasik et al. (2001)),
- number of residents in the household (Lindley et al. (1989)),
- postcode (Crook et al. (1992)),
- years at address (Crook et al. (1992)),
- income (Lindley et al. (1989), Crook et al. (1992), Banasik et al. (2001)),
- spouse's income (Banasik et al. (2001)),
- number of cards held (Carow and Staten (1999)),
- length of holding the card / years at bank (Lindley et al. (1989), White (1975), Crook et al. (1992)),

- the highest outstanding credit card debt (White (1975)),
- current outstanding balance (White (1975)),
- size of the transaction (White (1975)),
- outgoings (Banasik et al. (2001)).

The techniques used involved binary and multinomial logit, a probit regression. Banasik et al. (2001) looked at the 'desired' usage of the card holders as opposed to the observed usage, incorporating the credit-constrained aspect into the analysis and thus correcting for the sample selection bias. Only a few studies have investigated the time between purchases/ transactions. Till (2001) provided an investigation of the form of the distribution underlying the time between transactions for a store card. It was shown that the number of transactions up to time T can be modelled by a Negative Binomial, and that the times between transactions follow a Weibull distribution.

Ansell et al. (2001) examined the purchasing behaviour of the customers of an insurance company with the aim of deciding on the marketing strategy. The customers were segmented into defined groups, and the behaviour of these groups was explored. The major determinants were found to be age and financial sophistication. Younger and less sophisticated customers tended to make further purchases quickly or not to make them at all. The response from older and more sophisticated customers was more extended in time and it was not possible to determine the point beyond which they would not be expected to come back.

Thomas et al. (2003) also found age and financial status variables to be powerful predictors of the propensity to purchase financial products from an insurance company. In addition, economic variables (the consumer confidence index, changes in FTSE ALL Share index, House Price Index, average earning index and bank interest rate) had a strong and significant effect on the purchase behaviour. The study compared an analysis of multiple purchases to a separate purchase event analysis, and found that the magnitude and significance of most coefficients were similar, indicating relative insensitivity of results to segmentation on the number of purchases. This result justifies the decision to limit the scope of investigation of the current research to the second purchase only.

One of the most important issues is combining the results of estimating usage with the probability to default. Segmentation appears to be a standard approach. The common strategy for reviewing the credit limit is described in Thomas et al. (2002) when a measure of risk, a behavioural score, is split into a number bands and is tabulated against some measure of return, e.g. a monthly credit turnover, that is also categorised into bands. A credit limit is assigned to each cross-section between behavioural scores bands and credit turnover bands. It is noted though, that limits are chosen subjectively in the majority of cases. Banasik et al. (2001) showed that segmentation based on card usage improved the prediction of the customer's repayment performance. Oxley (2003) also reported a significant improvement in a portfolio's profitability from setting the credit limit by segmenting the customers on the basis of predicted usage and the likelihood of default, as compared to segmentation when usage was not taken into account.

6.3 Approach taken and data description

The data relates to a store card used in three different European countries, the same store card that was considered in previous chapters. The card is offered by a major international bank through a range of stores selling 'white' durable goods. The store card is normally taken out at the time of a first purchase under different agreement types: budget, deferred. Thus apart from selling some piece of merchandise at this initial point, the store receives the opportunity to develop a relationship with the customer with a view of further potential purchases. The lender is also interested in developing this relationship into profitable active usage of the card. In planning a retention strategy, it is necessary to know which individuals are going the make a subsequent purchase and when they are going to make it.

Figure 6.1 represents three routes that the customer can follow when s/he enters the system of revolving credit after making the first purchase:

1) go into delinquency, which will be manifested by increase in balance because of the missed monthly payments;
2) make the second purchase, which will also lead to balance increase;
3) repay the cost of the first purchased product without going into delinquency or making the second purchase, thus the balance will decrease or stay at 0.

It is assumed that the customer cannot make the purchase and miss the monthly payment simultaneously - there is at least one time period (month) separating these two events. The customer can leave the system either by closing the account after having repaid the cost of the purchase or by going into default so that the cost is eventually written off by the lender.

The interest of the current study focuses on the customer's propensity to take the two routes that will lead to the balance increase as represented in Figure 6.1. The previous chapter investigated the customer's propensity to default and looked at how quickly this can happen. This chapter investigates the propensity for further purchase. No desire to use the card further is difficult to observe. Whilst there is a special marker for closed accounts, only a small proportion of accounts have it (Table 5.2, 'Randomly right-censored' category). At the same time quite a significant number of accounts do not make the second purchase, although the account is not closed.

Figure 6.1 Potential Behaviour of the Customer

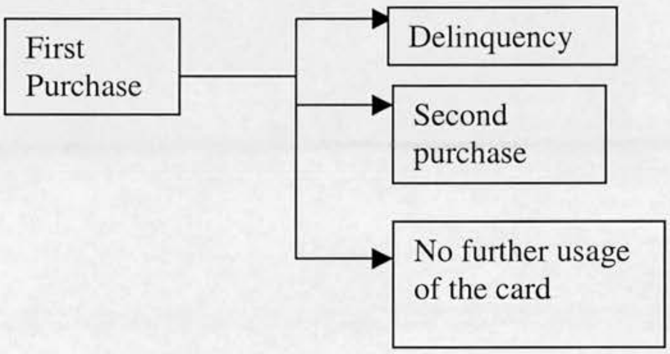
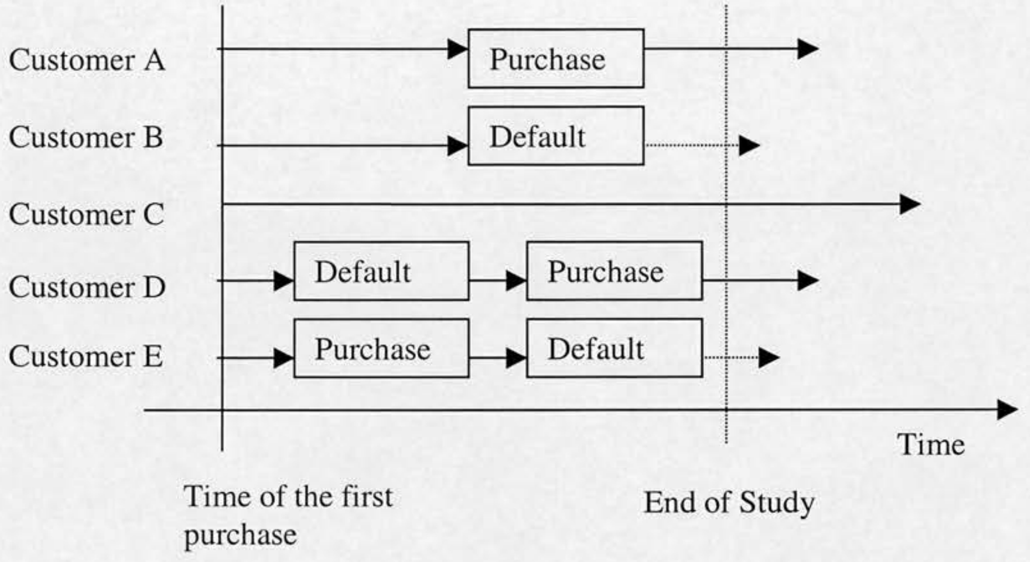


Figure 6.2 displays the behaviour of five typical customers. Customer A makes the second purchase within the study period and then continues on potentially to make further purchases. Customer B goes into delinquency and may or may not continue. Customer C does not make a further purchase within the study period. Customer C could either be one who will eventually make a second purchase or an individual who has no intention to make a second purchase. Customer D defaults (i.e. misses 2 consecutive payments, according to the definition of default adopted in previous chapters), but then makes repayments and so can make a further purchase before the end of the study period. Customer E makes a second purchase and defaults afterwards.

Figure 6.2. Different types of the observed behaviour



As a starting point, time to the second purchase is considered ignoring the impact of default. In this formulation all the customers that do not make a second purchase are censored at the end of the observation period. Another possible approach would be the competing risks model. However, this model assumes that the customer, after experiencing the event (increase in balance due to the purchase or missed payment), leaves the risk set, which is not the case in the context of the data. Under the competing risk formulation Customer D would be censored at the point of default and the subsequent purchase would be discarded from the analysis.

The time was measured from the point of the first purchase until the second purchase or until the end of the observation period. The period of observation ranged from 12 months to 25 months, the time to the second purchase could be from 2 months to 25 months.

Chapter 5 considered AFT and PH approaches within survival analysis. It was shown that AFT and PH models were very close in predictive accuracy, hence the choice of the model can be based on the additional benefits the approach can offer. The analysis in this chapter is based on the proportional hazards model because of its ability to handle time-dependent covariates. This feature is extremely useful in entering the behavioural data, as will be shown below. In order to retain only significant variables in the model, a stepwise selection procedure was applied.

The basic model to be estimated is

$$h(t)=h_0(t)e^{\beta' x},$$

where h_0 is an unspecified baseline hazard function, x is the vector of covariates and β is the vector of parameters that need to be estimated.

Information that can be used to predict the purchase can be grouped into 3 major categories:

- information about the applicant (personal data),
- information about the first intended purchase (purchase data),
- transactional information after the credit has been granted (behavioural data).

It is of interest to investigate the value of each information level in predicting time to the next purchase. A list of the characteristics used is presented in Table 6.1. Their values/levels were coarse-classified according to similarity in p_j where p_j denotes the probability that those cases within a coarse category, j, make a second purchase, and transformed into binary dummy variables, see Thomas et al. (2002).

Table 6.1 Variables used in the analysis

Information level	Characteristics available
Personal	Home telephone
	Residential status
	Marital status
	Occupation (Full-time, part-time, self-employed, etc.)
	Age
	Time at address since 18 years old
	Time in employment
	Type of business (Manufacturing, banking, catering, etc.)
	Employer's phone
	Spouse age
	Number of dependants
Initial Purchase	Product code
	Product price
	Payment date
	Card insurance
	Credit insurance
	Agreement type
Behavioural	Difference between the outstanding balance and credit limit at period t_i
	Delinquency status at period t_i
	Percent of outstanding balance repaid at period t_i

So for each of the three countries the proportional hazard models were fitted using all levels of information. The added value of each level of information was measured by the difference in Log Likelihood, which shows how much variation in data is accounted for by the model. The improvement in the model predictive power was measured by AUROC and error rate.

The transactional behavioural data changes every time period and two related ways of entering it into the model were examined:

1. by incorporating all information available for 25 months,

$$h(t) = e^{\beta' x + \gamma' z_{t-1}} h_0(t),$$

where z is a vector of time-varying behavioural variables and x is a vector of static application variables.

2. by entering the information which is only available at a certain period of time, an approach suggested by Stepanova (2001), Stepanova and Thomas (2001). The model which was estimated can be written as

$$h^s(t) = e^{\beta'(s)x + \gamma'(s)z_{t-1}} h_0^s(t),$$

where s indicates the number of time periods elapsed since the first purchase.

With both approaches the parameters for time-dependent variables are re-estimated at each period of time. But the first approach produces parameter estimates aggregated over the whole period of observation, thus uses the most complete information available. The second approach explicitly breaks down the partial likelihood estimation for each time period, tailoring the parameters to a given point in time, and although it uses a limited amount of historical information, this information is the most recent and probably, most relevant information.

We investigated the value of historical behavioural information. The initial assumption was that all information needed to predict the next purchase is the most recent one. However, inclusion of lagged variables improved the model fit (see Section 6.4.3.3), and it is likely that adding more lagged variables would improve the model fit further. Nevertheless, the decision was taken not to continue the exploration in this direction, since it was shown by Stepanova (2001) that incorporating all behavioural information accumulated to date resulted in high collinearity among the model parameters with a very modest increase in predictive accuracy.

6.4 National patterns in describing time to the next purchase

Table 6.2 provides the size of the samples used in the analysis. Each dataset was randomly split into 70% (training sample) and 30% (hold-out). The hold-out sample was reserved for testing the predictive ability of the models.

Table 6.2 shows that the most active buyers are the Dutch card-holders, and Germans are the least active ones. There is also a notable difference in the percentage of censored observations between good and bad payers for all three countries.

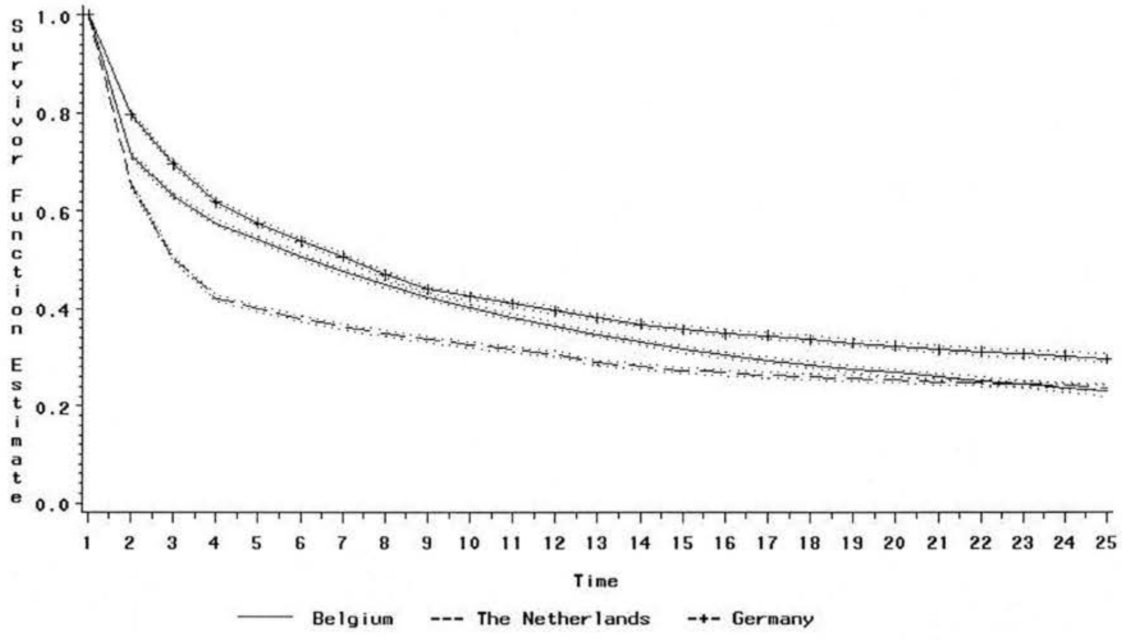
Table 6.2 Samples used in modelling time to the second purchase.

Performance	**Total**	**2-nd purchase**	**Censored**	**%Censored**
		Belgium		
Good	22708	16648	6060	26.69%
Bad	3084	1634	1450	47.02%
Total	25792	18282	7510	29.12%
		The Netherlands		
Good	24520	19515	5005	20.41%
Bad	7589	4141	3448	45.43%
Total	32109	23656	8453	26.33%
		Germany		
Good	66939	45217	21722	32.45%
Bad	8909	4529	4380	49.16%
Total	75848	49746	26102	34.41%

The samples selected in this Chapter do not match exactly the samples used previously. Some agreement types in the Dutch dataset showed very low chances of making the second purchase. It was decided to remove these observations from the analysis, because most likely it was the condition of the agreement that prevented the second purchase. Further, it was decided to limit the analysis of the time to the second purchase to agreement types common to all three countries. Otherwise the samples were as used in Chapter 4.

The baseline survival curve in Figure 6.3 shows that 50% of the customers have made a second purchase within 3 months in the Netherlands, within 6 months in Belgium and within 7 months in Germany. Approximately one third of the card-holders do not make a second purchase within 25 months. The curve will be slightly affected by the accounts that were not observed for the whole period of 25 months. Censored customers can be considered as a lost opportunity, since potentially these customers could have made more purchases within the time period observed.

Figure 6.3 Baseline SDF with 95% CI, personal data



6.4.1 Personal data

Personal characteristics are normally used by the lender to estimate the risk of default and thus to make a decision whether to accept an applicant or not. It is of interest to see whether this information can provide any insights into the likelihood of repeated purchase. The parameter estimates from applying stepwise proportional hazards using just the personal data are in Table 6.3 column 1. The hazard ratios (exponentiated parameters, sometimes referred to as Relative Risk or Risk Ratio) are given in Appendix, A25.

There are notable similarities and also differences between countries. 'Home Phone Given' indicates higher chances of making the second purchase in Germany, but it is not significant in the other two countries. Number of children is insignificant for Belgium, in Germany families with a greater number of kids are more active buyers, and in the Netherlands the customers that make the second purchase fastest are those that do not provide any information about their children.

Table 6.3 Parameter estimates for models with different levels of information. Belgium

Variable	Model								
	Personal	Personal, Purchase	Personal, Purchase, ATS_1	Personal, Purchase, ATS_{t-1}	Personal, Purchase, ATS_{t-1}, ATS_{t-2}	From 2nd period- more info	Personal, Purchase, ATS_{t-1} – Goods	Personal, Purchase, ATS_{t-1} -Bads after 2nd p	Personal, Purchase, ATS_{t-1} – Bads before 2 p
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Phone number given								0.264	0.615
Kids: 1 child									0.329
Marital: Widowed	0.111	0.118	0.135						
Res: Renting room, parents	0.086	0.079	0.086		0.104	0.119	0.114		
Res: Renting house/flat	0.104	0.084	0.096	0.081	0.137	0.161	0.142		
Occup: Retired		-0.078							
Occup: Part-time	0.105								
Occup: Self-employed	-0.226	-0.122	-0.102	-0.126	-0.133			-0.465	
Type of business – Unknown	0.139	0.089			0.121				
Type of business – 21	0.188	0.115	0.082	0.109	0.147	0.117	0.077		
Age : under 21		-0.129							
Age : 22-27						0.069			
No spouse		-0.077	-0.069		-0.059			-0.197	
Time at address : 6 months	0.061		0.062		0.119	0.136	0.086	-0.136	
Time at address : 6 mths – 1 yr	0.104	0.052	0.105	0.045	0.151	0.166	0.116		
Time on job: 1 yr						0.062			
Allowance			0.120	0.088		0.175	0.136		
No card insurance		-0.076	-0.078	-0.061	-0.107	-0.124	-0.083		
No credit insurance		0.121	0.088		0.128	0.128	0.106		

Table 6.3. Parameter estimates for models with different levels of information. Belgium (continued from previous page)

	1	2	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Product type- computers		-0.118		-0.117	-0.217	-0.201		0.280	
Product type- TV					-0.061				
Product type- household1									0.364
Product type- household2									0.369
Product type- phones		0.056	0.047	0.055		0.073	0.090		
Price = 0		0.280		0.149	0.589	0.483			
Price < =10,000 BEF		0.630	0.208	0.487	0.706	0.695	0.261		0.341
10,000 < Price <= 16,000 BEF		0.487	0.085	0.367	0.529	0.471	0.110		
16,000 < Price <= 20,000 BEF		0.328		0.224	0.316	0.274			
20,000 < Price <= 40,000 BEF		0.205		0.153	0.246	0.227	0.058		
Pay date 01		-0.319	-0.314	-0.300			-0.328	-0.376	
Pay date 08		0.169	0.137	0.158	0.334	0.328	0.133		-0.407
Pay date 14,15		0.372	0.351	0.360	0.221	0.206	0.363	0.284	
Agreement type – budget		0.726	0.854	0.775	0.452	0.675	0.878	1.209	
ATS $_{t-1}$ – over credit limit			-1.095	-3.867	-4.607	-2.540	-1.107	-1.251	
ATS $_{t-1}$ = 0			-0.235	0.574	0.624	1.223	-0.171	-0.631	
ATS $_{t-1}$ <= 5000 BEF			-0.423	-0.318	-1.156	-0.451	-0.359	-0.554	
5000< ATS $_{t-1}$ < =10,000 BEF			-0.222		-0.315		-0.194	-0.305	
10,000< ATS $_{t-1}$ <= 20,000 BEF			-0.128				-0.140		0.282
ATS $_{t-2}$ – over credit limit					0.878				
ATS $_{t-2}$ = 0					0.956				
ATS $_{t-2}$ <= 5000 BEF					0.517				
5000< ATS $_{t-2}$ < =10,000 BEF					0.142				
1 missed payment						-4.855			
2 missed payment						-4.339			

Table 6.3. Parameter estimates for models with different levels of information. The Netherlands.

Variable	Model								
	Personal	Personal, Purchase	Personal, Purchase,ATS$_1$	Personal, Purchase, ATS $_{t-1}$	Personal, Purchase, ATS $_{t-1}$, ATS $_{t-2}$	From 2nd period- more info	Personal, Purchase, ATS $_{t-1}$ – Goods	Personal, Purchase, ATS $_{t-1}$ -Bads after 2nd p	Personal, Purchase, ATS $_{t-1}$ – Bads before 2 p
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Phone given		0.099	0.089	0.066	0.183				0.304
Kids - no info	0.253								
Marital: Married					0.085				
Marital: Single, Widowed		-0.058	-0.060	-0.047	-0.102	-0.120			
Res: Rented house, flat	0.179								
Res: Rented room	0.103	0.065	0.068						
Res: Parents	0.092	0.112	0.093	0.091	0.097	0.070	0.062	0.118	
Occup: Self-emp	-0.208	-0.216	-0.159	-0.163	-0.158				
Occup: Full-time		-0.067							
Type of business: Catering, shop	0.089	0.065	0.065	0.057			0.052		
Type of business: Benefit, agency	0.124								
Age: under 22		-0.229	-0.223	-0.182	-0.211				
Age: 22 - 31	-0.064	-0.143	-0.147	-0.134	-0.073				
Age: 32-36		-0.106	-0.095	-0.089					
Age: 37 - 49		-0.077	-0.060	-0.053	-0.084				
Spouse age: 27-34						-0.080	-0.053		
Spouse age: 35-47	0.069	0.050						0.205	0.269
Time address: up to 4 m	0.122	0.069	0.072	0.070		0.070			
Time address: 1y - 5y	-0.034							-0.132	
Time on job: 1.5y -3y		0.038	0.050	0.049					
Time on job: 3y -10y	-0.063							-0.101	

Table 6.3. Parameter estimates for models with different levels of information. The Netherlands (continued from previous page)

	1	2	*3*	*4*	*5*	*6*	*7*	*8*	*9*
No card insurance		0.383	0.366	0.386	0.427	0.480	0.421	0.260	
No credit insurance		0.183	0.164	0.181	0.244	0.280	0.269		
Credit ins 8		0.231	0.214	0.227	0.313	0.370	0.294		
Product: computers		0.212	0.201	0.216	0.279	0.330	0.308		
Product: video		0.275	0.254	0.265	0.335	0.310	0.271	0.248	
Product: TV		0.382	0.304	0.392	0.515	0.440	0.370	0.621	0.498
Product: furniture		0.263	0.135	0.194	0.321	0.240	0.180	0.411	
Product: cycles		0.219	0.099	0.137	0.205	0.160	0.117	0.405	
Price: under 400 NLG		0.119			0.147			0.384	
Price: 401-800 NLG					0.094	0.090		0.297	
Price: 801-1600 NLG		-0.262	-0.222	-0.235	-0.155	-0.120	-0.207	-0.517	-0.349
Price: 1601-2200 NLG		0.084			0.067			0.275	
Price: 2201-3000 NLG								0.234	0.273
Agreement*pay date: budget*01		1.305	1.335	1.357	1.231	1.390	1.195	1.594	0.640
Agreement*pay date: budget*other					-0.239	-0.390	-0.321	0.675	-0.715
Agreement*pay date: defer		1.011	0.989	1.029	1.074	0.950	0.923	1.296	1.162
ATS $_{t-1}$: over credit limit			-0.351	-0.794	-1.759	-0.250	-0.221	-1.387	-2.565
ATS $_{t-1}$: 0-500 NLG			-0.181			0.300	0.201	-0.438	-1.106
ATS $_{t-1}$: 501-1000 NLG				0.117		0.440	0.309	-0.350	
ATS $_{t-1}$: 1000-3000 NLG			0.045	0.073	-0.126	0.340	0.241	-0.301	
ATS $_{t-2}$: over credit limit					1.064				
ATS $_{t-2}$: 0-500 NLG					0.960				
ATS $_{t-2}$: 501-1000 NLG					0.589				
ATS $_{t-2}$: 1000-3000 NLG					0.801				
1 missed payment						-3.200			
2+ missed payments						-3.830			
Percent repaid						0.270			

Table 6.3. Parameter estimates for models with different levels of information. Germany

Variable	Model								
	Per-sonal	Personal, Purchase	Personal, Purchase, ATS_1	Personal, Purchase, ATS_{t-1}	Personal, Purchase, ATS_{t-1}, ATS_{t-2}	From 2[nd] period-more info	Personal, Purchase, ATS_{t-1} – Goods	Personal, Purchase, ATS_{t-1} -Bads after 2[nd] p	Personal, Purchase, ATS_{t-1} – Bads before 2 p
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Phone given	0.096	0.136	0.118	0.108	0.150			0.208	
Kids: 1-2		0.054	0.041	0.031	0.058	0.039	0.027		
Kids: 3+	0.091	0.150	0.142	0.134	0.148	0.132	0.136	0.161	
Marital: Widowed	0.110								
Res: Owner	-0.039								
Res: Rented house	0.145	0.145	0.139	0.137	0.152	0.148	0.133		
Occup: Full-time	-0.139							0.229	
Occup: Self-emp	-0.210	-0.060			-0.067	-0.072			
Type of business: Service	-0.072								
Type of business: Education, healthcare	-0.184	-0.045	-0.044	-0.044	-0.051	-0.038	-0.041		0.240
Age: under 21	0.064	-0.073					0.114		
Age: 21 - 29	0.097		0.031	0.030	0.056	0.039	0.072		
Age: 30-34	0.071				0.040		0.035		
Age: 40-55			-0.030	-0.030		-0.047			
Age: 55+		-0.075	-0.138	-0.144	-0.066	-0.138	-0.139		
Spouse age: 40-62	-0.064								0.484
Time address: up to 6m		-0.044	-0.036		-0.059				
Time address: 6m - 2y	0.031								
Time on job: missing			0.072	0.077			0.091		
Time on job: up to 2y			0.037	0.042		0.049	0.088		

Table 6.3. Parameter estimates for models with different levels of information. Germany (continued from previous page)

	1	2	3	4	5	6	7	8	9
Time on job: 2-4.5y							0.042		
Time on job: 4.5y -7.5y	-0.033								
Time on job: 7.5y -10y	-0.057								
No card insurance		-0.217	-0.228	-0.238	-0.155	-0.198	-0.278	-0.136	
Price: under 750 DEM		0.383	0.295	0.301	0.347	0.344	0.345		-0.353
Price: 751-1300 DEM		0.278	0.201	0.209	0.245	0.249	0.242		-0.286
Price: 1301-1800 DEM		0.163	0.113	0.124	0.147	0.144	0.133		
Price: 1801-2300 DEM		0.076	0.046	0.057	0.072	0.069	0.059		
Product: kitchen		-0.133	-0.135	-0.140	-0.149	-0.162	-0.180		
Product: computers		-0.053	-0.056	-0.063	-0.079	-0.080	-0.099		-0.187
Product: household							-0.070		
Product: video		0.059	0.052	0.050				0.203	0.305
Pay date: 01		1.386	1.391	1.396	1.552	1.640	1.517	0.884	0.950
Pay date: 08		0.143	0.146	0.141			0.121	0.300	
Agreement: budget		-0.216	-0.190	-0.192	-0.330	-0.248	-0.136	-0.481	-2.935
ATS $_{t-1}$: over credit limit			-0.473	-2.031	-4.156	-1.700	-1.132	-2.987	-2.744
ATS $_{t-1}$: 0-250 DEM			-0.349	-0.569	-1.160	-0.329	-0.431	-0.898	-1.591
ATS $_{t-1}$: 251-500 DEM			-0.106	-0.179	-0.500		-0.135	-0.246	-0.506
ATS $_{t-1}$: 501-1500 DEM						0.068		-0.091	
ATS $_{t-2}$: over credit limit					0.329				
ATS $_{t-2}$: 0-250 DEM					0.590				
ATS $_{t-2}$: 251-500 DEM					0.172				
1 missed payment						-4.124			
2+ missed payments						-4.602			
Percent repaid						0.942			

Whilst for Belgium and Germany widows are more active spenders on the card, marital status variables are not significant in the Netherlands. For all three countries those in rented accommodation and only recently at address are more likely to make a further purchase on the card. Self-employed people seem to be poor buyers across the three countries. Type of business appears to be significant in Belgium, but since there is no explanation behind the code used by the lender, it is impossible to match the categories selected in the Belgian model with other countries. For the Netherlands those working in catering and retailing are more active users than those that receive benefit or work through an agency. And in Germany education, healthcare and service workers are less likely to make the second purchase than the other professional groups.

Applicant's Age, which is normally found highly significant in estimating default (Crook (1997)), does not enter the Belgian model. In the Netherlands people between 22 and 30 are less likely to make the second purchase, while in Germany the chances of making the second purchase first, increase with age, but after 30 start to go down. Longer years on job are associated with less likelihood of the purchase for Germany and the Netherlands, whereas in Belgium Time on Job is not selected into the model.

Table 6.4 presents the application variables ranked in the order of significance, when several variables of one characteristic entered the model, only the most significant one was selected in order to simplify the comparison. It is interesting to note that top rows are taken by occupation, type of business and residential status.

However, the variation accounted for by the application model is modest. As can be seen from Table 6.6 the difference in Log Likelihood between the application model and no covariates model is only from 99 (Belgium) to 224 (Germany). Looking at the prediction results (Table 6.8) one cannot say that personal data is a powerful predictor. With AUROC ranging between 0.553 (Belgium) and 0.571 (Germany) the performance of the personal data models is only marginally better than the random classification into two classes. So additional information is needed in order to make better decisions on the retention of customers.

Table 6.4 Personal characteristics ranked in order of importance

Belgium	The Netherlands	Germany
Occupation: Self-employed	Number of kids - no info	Occupation: Self-employed
Type of business - 21	Occupation: Self-employed	Type of business: Service
Marital status-Widowed	Residential: Renting house/flat	Residential: Renting house/flat
Time at address- 6m-1y	Type of business - benefit, agency	Marital status-Widowed
Residential: Renting house/flat	Spouse age: 35-47	Age: 21 - 29
	Age: 22 - 31	Home phone given
	Time on job: 3y -10y	Kids: 3+
		Spouse age: 40-62
		Time on job: 7.5y -10y
		Time at address: 6m - 2y

6.4.2 Purchase data

Information about the first purchase made or about the intended first purchase expands the scope of the analysis to include such characteristics as the nature of the initial product, its price, agreement type, date of payment and whether the customer is taking out insurance on either the card or credit. The results of fitting both personal and purchase data together using the stepwise model are presented in Table 6.3 and Appendix, A25, column 2.

It is notable that the variables previously included appear again with only minor modification in the parameter estimates, suggesting there are no apparent problems with collinearity. For Belgium, the main changes are the introduction of Age and Spouse's Age, and the downgrading of Time at Address. Younger individuals are more likely not to make a second purchase as are those without a spouse, and hence the effect on Time at Address is seen.

For the Netherlands, Marital Status appears, and single/ widowed group shows lower chances to make a purchase (unlike Belgium and Germany). More age categories enter that show the increasing tendency for purchase with increasing age.

In Germany some application variables leave the model (e.g., Marital Status), indicating that this information is replaced by the information contained in the purchase variables. There are changes in the age categories selected, but the trend remains unchanged: younger and older customers are poorer buyers compared to middle-aged customers.

Nearly all purchase variables enter. Not taking card insurance seems to be an indicator of less likelihood of further purchase for the three countries. No credit insurance implies a greater chance of purchase for Belgium and the Netherlands, but is not selected by the German model. Product type is highly significant. Those buying computers seem to have relatively lower potential for a subsequent purchase. In Belgium phone buyers exhibit the higher chances for the second purchase, whereas in Germany and the Netherlands the higher likelihood for the repeated purchase is associated with the initial purchase of a TV or video. Those who buy lower cost items initially are more likely to make the second purchase on the card. This may be explained by the fact that they need less time to repay the cost of a product.

The payment date has a varying effect. In Belgium those making payments later in the month are more likely to use the card again, but Germany shows the opposite trend. For agreement type those using budget plans in Belgium are much more likely to make further purchases on the card, but less likely to do so in Germany. In the Netherlands agreement type was found to be strongly associated with payment date, that is why, the two characteristics were combined together. Those with budget agreement making payments on the 1st day of the month have the highest chances for a repeated purchase, chances of deferred agreement schemes are slightly less, but still higher if compared to other agreement types.

But for all three countries agreement type, payment day and product price are the most powerful discriminators that outweigh the application variables in significance (Table 6.5). The model fitted accounts for considerably more variation than the previous model. The range of improvement (Table 6.6) is between 979 (Belgium) and 7097 (Germany). This translates into quite considerable increases in predictive accuracy (Table 6.8a) from 0.553 to 0.689 (Belgium) and from 0.571 to 0.812 (the Netherlands).

Table 6.5 Personal and purchase variables ranked in order of importance

Belgium	The Netherlands	Germany
Agreement type – budget	Agreement*pay date: budget*01	Pay date 01
Price < =10,000 BF	Product: computers	Price: under 750
Pay date 14,15	Price: under 400 NLG	No card insurance
Age : under 21	No card insurance	Agreement type – budget
Self-employed	Age: under 22	Kids: 3+
No credit insurance	Occupation: Self-employed	Residential: Renting house/flat
Marital status-Widowed	Residential: parents	Home phone given
Product type- computers	Home phone given	Product: kitchen
Type of business – 21	No credit insurance	Age: 55+
Renting house/flat	Time at address : up to 4 mths	Occupation: Self-employed
No spouse	Type of business - catering, shop	Type of business: Service
No card insurance	Marital status: Single, widowed	Time at address: 6m - 2y
Time at address : 6 mths – 1 yr	Spouse age: 35-47	
Time on job: 1 yr	Time on job: 1.5 - 3y	

6.4.3 Transactional information

6.4.3.1 Credit availability

After the credit limit is assigned and the card is issued, the transactions are recorded on a monthly basis. It is of interest to see how much behavioural measures add to the predictive accuracy of the models.

Variables representing the customer's behaviour vary with time and therefore there is a need to decide how they should be fitted into the model. At the point of initial purchase it is possible to calculate the difference between the outstanding balance and the credit limit. This variable is called 'amount to spend', (ATS), and represents the credit availability to the customer when s/he first enters the credit agreement.

There are several modelling options. First, one can use ATS at the time of the first purchase, period 1. This value of ATS is denoted as ATS_1. The results when using this variable are presented in Table 6.3, column 3.

However, it seems more likely that the *most recent* ATS will provide a better explanation for an individual's second purchase. Hence the second model contemplated was to include the ATS for the period preceding the second purchase into the model as a time-varying variable, ATS_{t-1}. Here one can use the training sample for the whole period of observation to estimate the parameters of the static application variables and time-varying ATS_{t-1}. This means that at each stage of partial likelihood estimation a different value of ATS will be entered into the model and that will be done for all 25 periods. The results of fitting this second model are given in Table 6.3, column 4.

Introduction of either ATS_1 or ATS_{t-1} results in a small number of variables leaving the model, e.g. Applicant's Age for Belgium, Spouse's Age for the Netherlands and Occupation for Germany. But overall, the parameters remain unchanged. For all three countries ATS variables are significant and have a major effect. As expected, those with higher values of ATS are more likely to make the second purchase.

The reduction in the magnitude of Log Likelihood (Table 6.6) indicates that most recent ATS is superior to the ATS at the point of first purchase. For ATS_1 the reduction as compared to the 'Personal + Purchase' model ranges from 94 (the Netherlands) to 298 (Belgium). Time-varying ATS_{t-1} leads to a reduction of 478 (the Netherlands) to 1508 (Germany), emphasising the importance of recent ATS over initial ATS.

Looking at all models that incorporate the information sequentially, as it arises, it is possible to conclude that for Belgium the most marked improvement in the model fit is observed when the first purchase information is added and when the most recent ATS enters the model. For the Netherlands and Germany the improvement is the most notable one with purchase characteristics available. Although ATS also decreases the magnitude of Log Likelihood, it cannot really match the scale of improvement obtained from knowledge about the characteristics of the initial purchase.

Table 6.6 Log Likelihood statistics for models with different levels of information

Model	Log L	df	Log L difference[12]
Belgium			
No covariates	-120280		
Personal data	-120181	9	99
Personal+purchase	-119202	23	979
Personal+purchase +ATS_1	-118904	23	298
Personal+purchase +ATS_{t-1}	-117999	20	905
The Netherlands			
No covariates	-161015		
Personal data	-160873	12	142
Personal+purchase	-157876	27	2997
Personal+purchase +ATS_1	-157782	26	94
Personal+purchase +ATS_{t-1}	-157304	25	478
Germany			
No covariates	-369628		
Personal data	-369404	16	224
Personal+purchase	-362307	20	7097
Personal+purchase +ATS_1	-362106	25	201
Personal+purchase +ATS_{t-1}	-360598	24	1508

An alternative approach to fitting the model with time-varying ATS_{t-1} over 25 months would be to re-estimate the model at each time period, as was suggested by Stepanova (2001) and Stepanova and Thomas (2001). Thus for every time period *t* the model is fitted to the training risk set which only includes accounts that have survived up to this period. The approach has an advantage of giving a snapshot of parameter estimates at each month, so it is possible to see the dynamics of the parameter changes. However, since, as the baseline survival curves (Figure 6.3) show, the second purchase is made fairly quickly, the risk set will be shrinking rapidly, and after 12 months there will be 40% or even less remaining of the original sample. So the decision was made to estimate the parameters for the first 10 months, the period of the greatest interest, because the majority of events take place within this period.

[12] The difference is taken between the models in 'adjacent' rows, e.g., LL difference for 'personal+purchase' model is the LL taken from the LL of the 'personal' model

The score (**$\beta'x$**) from the 'Personal+Purchase' model was used at each time period instead of the list of personal and purchase variables. This has the limitation of loosing any time-varying patterns that may exist between behavioural variables that were allowed to change with time and personal and purchase variables that were entered into the model as a single 'application score'. But the gain was a much higher speed of estimation and, besides, the main objective was to look at how parameters of behavioural variables change over time.

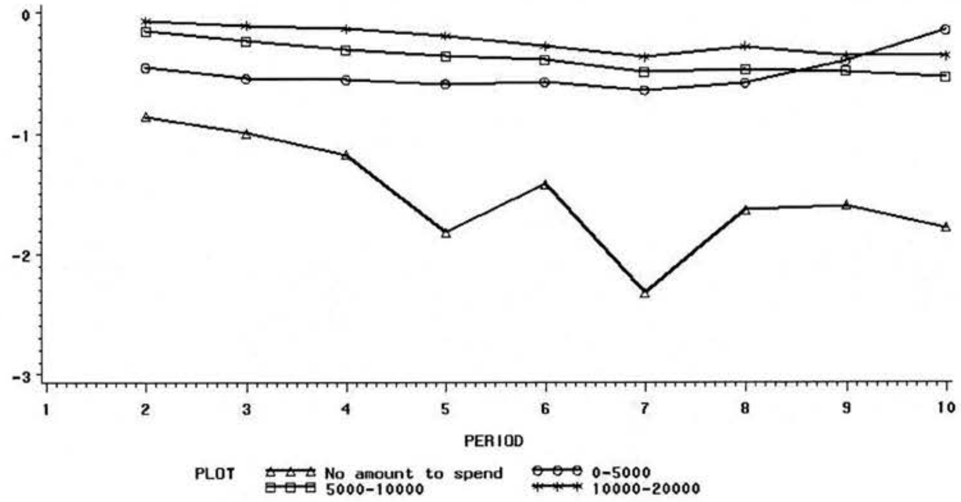
For Belgium, the initial coarse-classification of ATS involved 5 categories. Since the number of people falling into the category of 'No amount to spend' was rapidly going down with time, it was decided to group this category with 'Over credit limit'. For the Netherlands and Germany ATS was split into 4 categories in both the 'total observation period' model and the 'specific time period' model.

As can be seen from Figure 6.4 in all three countries two categories with the higher values of ATS (groups 3 and 4) demonstrate a roughly proportional effect at each time point. But in Belgium and the Netherlands these groups deviate from 0, indicating the increasing importance with time and decreasing chances for the second purchase. In Germany these two groups converge to 0. Group 2, the group with the smaller ATS values, stays proportional to groups 3 and 4 most of the time in Belgium and Germany, but closer to period 10 crosses these other two lines. This can be an indication of increased variability because of the shrinking risk set. Group1 'Over credit limit' shows a very different behaviour, these accounts are becoming less and less likely to make the 2nd purchase.

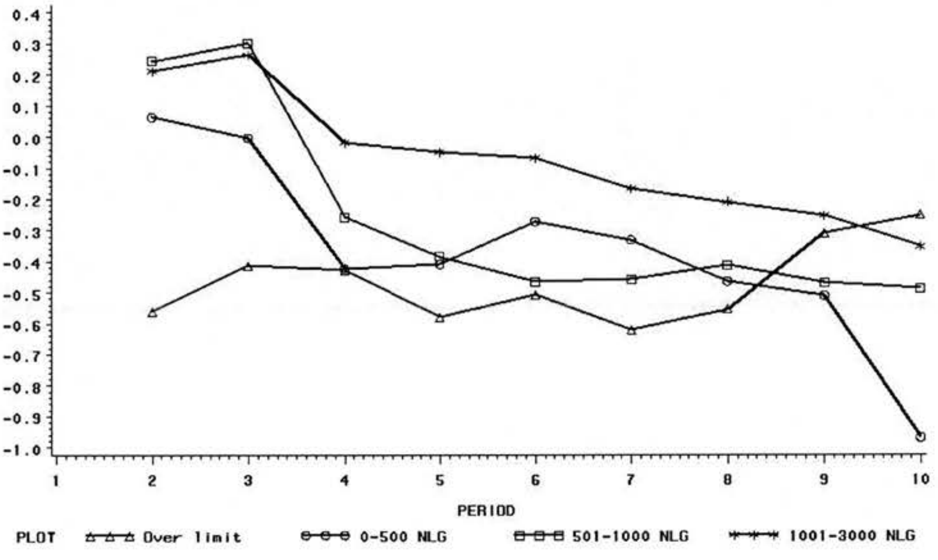
In the Netherlands Group 2 deviates from 0, but from month 4 to 9 it stays relatively flat, indicating that accounts with this ATS have a relatively time-invariant propensity for 2nd purchase for this period, that goes sharply down at month 10. It is interesting to note that 'Over credit limit' fluctuates between -0.4 and -0.6 up to month 8 then goes closer to 0. This indicates that accounts that spend beyond their limits have a relatively unchanging probability of making the 2nd purchase during the first 8 months.

Figure 6.4 Parameter estimates for Amount to Spend within 10 months.

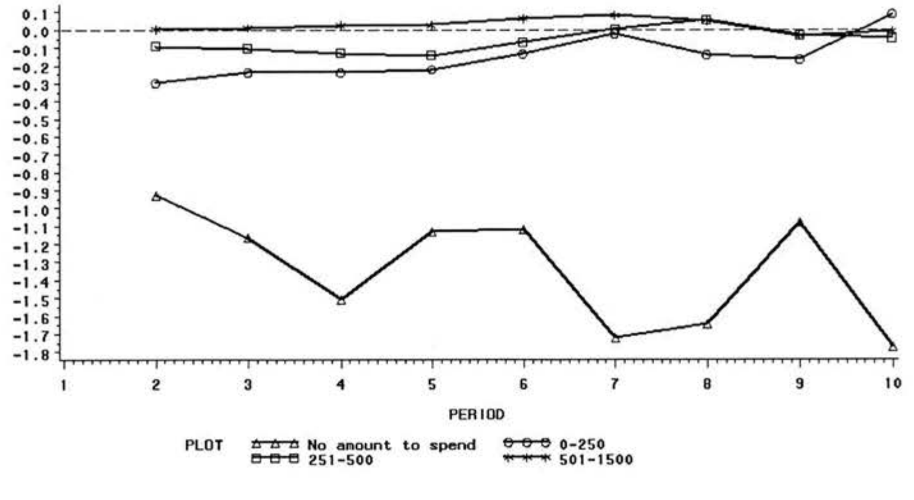
Belgium



The Netherlands



Germany



6.4.3.2 Delinquency status and repayment dynamics

Delinquency status was recorded on a monthly basis, starting from the month of the initial purchase until the end of the observation period. So starting from period 2 it was possible to use the delinquency status from the previous period. It was also possible to observe the repayments dynamics, by looking at how much of the outstanding balance was repaid in the previous month. For each period of time ($t \geq 2$) three more variables were added: an indicator of 1 missed payment, an indicator of 2 or more missed payments (our definition of default), and percent repaid of the outstanding balance at the period preceding the purchase.

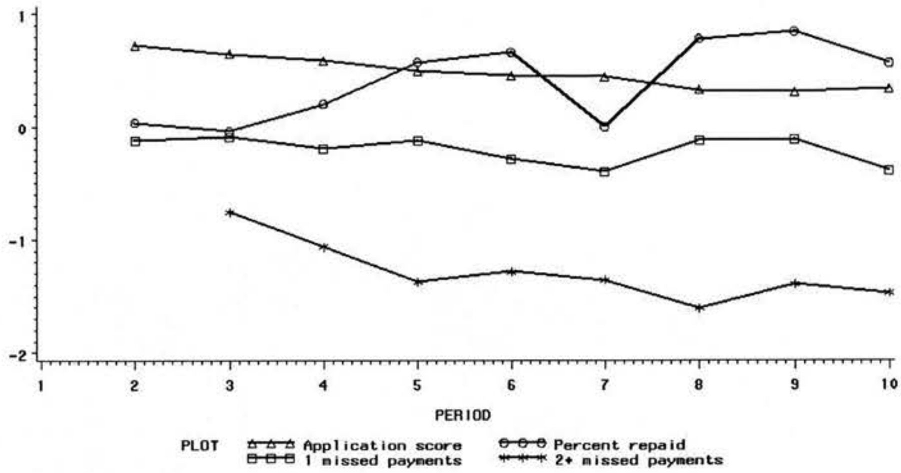
The model with time-dependent behavioural variables was fitted to the risk set of cardholders that make the purchase at period 3 and beyond. The parameter estimates and hazard ratios are shown in Table 6.3 and Appendix A25, column 6. Both delinquency status variables are highly significant for all three countries and have a major effect on the time to the 2nd purchase. Percent repaid was not selected into the model in Belgium, but turned significant in Germany and the Netherlands.

Figure 6.5 shows that the parameter estimates of application score decrease with time for all three countries, although the character of the decrease is different. For Belgium it happens gradually, for the Netherlands it drops abruptly between periods 3 and 4 and then stays flat, for Germany it starts to decrease from period 7. This is expected since the information becomes more historic. This highlights the difference between this investigation and previous work on defaults on fixed terms loans, (see Stepanova (2001), Stepanova and Thomas (2001)).

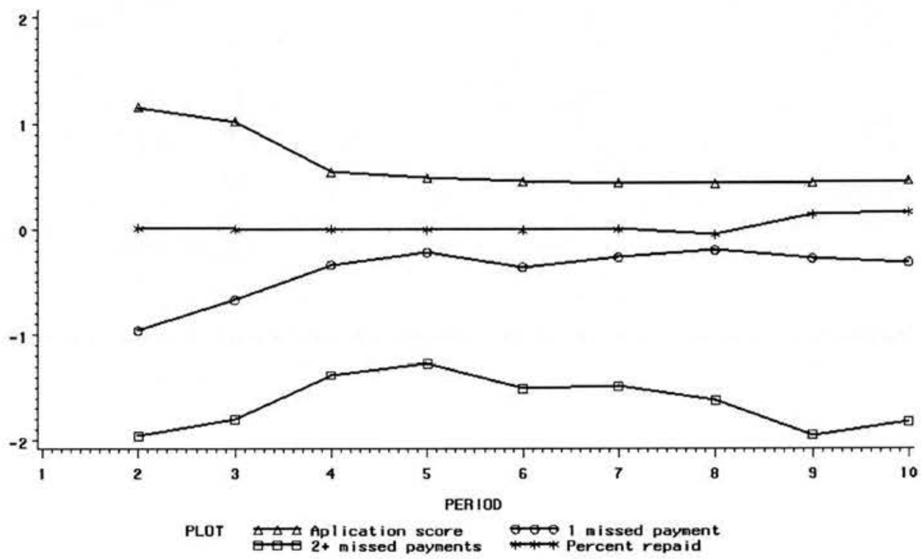
The behavioural aspects, on the contrary, become of greater importance over time. The only behavioural variable that does not entirely follow this general pattern is '1 missed payment' that gets closer to 0 for Germany and the Netherlands.

Figure 6.5 Parameter estimates for Application score, Delinquency, Percent repaid within 10 months

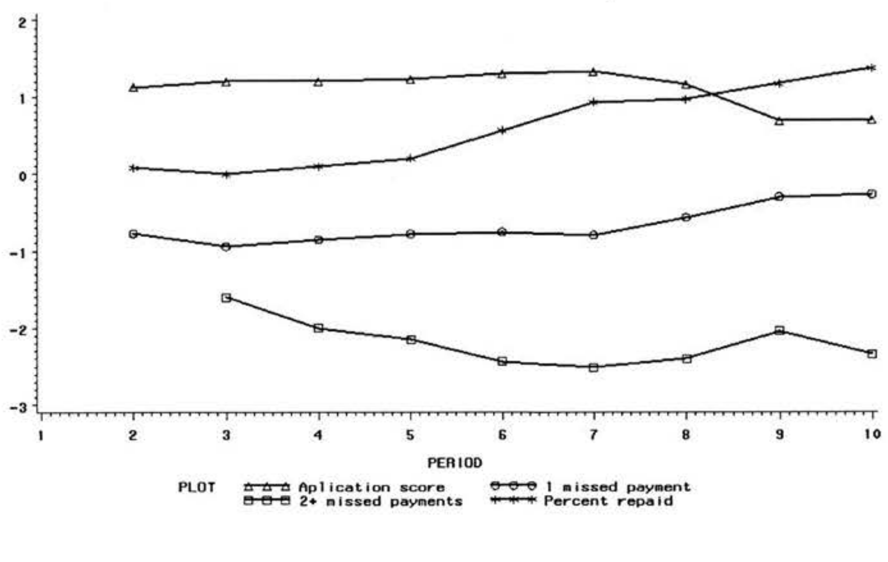
Belgium



The Netherlands



Germany



6.4.3.3 Testing for the Markov property

The analysis presented above rests on the assumption that it is the most recent behavioural information that is of greatest value. This is equivalent to saying that only the status at period preceding the purchase is important, and the information before this point is irrelevant.

An investigation of whether this assumption holds was undertaken for ATS. The model was fitted with ATS two periods before the purchase and both ATS one and two periods before the purchase. Table 6.7 compares the Log Likelihood statistics for the model with ATS_{t-1} and ATS_{t-2} applied to the sample of cardholders that make the second purchase starting from period 3.

Table 6.7 Log Likelihood statistics for the ATS lagged and not lagged.

Model	Log L	df
Belgium		
No covariates	-72778	
Personal+purchase +ATS_{t-1}	-71590	23
Personal+purchase +ATS_{t-2}	-71678	24
Personal+purchase +ATS_{t-1}+ATS_{t-2}	-71183	22
The Netherlands		
No covariates	-75997	
Personal+purchase +ATS_{t-1}	-79973	26
Personal+purchase +ATS_{t-2}	-80086	25
Personal+purchase +ATS_{t-1}+ATS_{t-2}	-79822	29
Germany		
No covariates	-253266	
Personal+purchase +ATS_{t-1}	-245524	22
Personal+purchase +ATS_{t-2}	-246039	22
Personal+purchase +ATS_{t-1}+ATS_{t-2}	-245479	25

The inclusion of ATS_{t-1} and ATS_{t-2} results in quite notable increases in the amount of variation accounted for, and both variables are significant. This suggests that ATS has a more complicated behaviour than a first order Markov chain.

One possible solution of incorporating the behavioural information from all time periods preceding the event is suggested in Stepanova (2001). The model is re-estimated for each period of time, but each time the application score is replaced by the resulting score from the previous period model. In this way all preceding information is accumulated in one variable. However, it was found that this resulted in high correlation between variables, and therefore an unstable performance. Therefore, the inclusion of historic information was not pursued further.

6.5 Predictive performance of national models

The predictive ability of the models was tested on hold-out samples taken from risk sets for periods 1 to 6. This was done in order to see how the performance of different models changes with time. At month 1 (initial purchase) we had a full risk set, since no one could have made the second purchase yet. This risk set, R_1, was split randomly into a development (70%) sample, D_1, and a hold-out sample, H_1. The models ('Personal', 'Personal+purchase', 'Personal+purchase + ATS_1', 'Personal + purchase +ATS_{t-1}' and 'Application score +ATS(s)' for s=1) were developed on D_1 to predict time to the second purchase, $T \geq 2$, and tested on H_1. The results are given in Table 6.8. The number in brackets gives the number of observations in the hold-out sample.

Then at month 2 a risk set, R_2, contained only accounts that did not make a second purchase in this month, those accounts that did make a purchase were removed from the risk set. The risk set, R_2, was split randomly into a development (70%) sample, D_2, and a hold-out sample, H_2. The development sample, D_2, was used to build the models 'Personal + purchase + Z_{t-1}' and 'Application score +ATS(s)', 'Application score + $Z(s)$' for s=2 to predict time to the second purchase, $T \geq 3$, with Z denoting all behavioural characteristics. These models together with those developed on D_1 ('Personal', 'Personal+purchase', 'Personal+purchase + ATS_1', 'Personal + purchase +ATS_{t-1}') were tested on H_2. The cut-offs were set so that the number of predicted purchases was equal to the number of actual purchases in the corresponding hold-out sample.

The same procedure continued up to month 6. We decided to stop after period 6 because, firstly, the size of the hold-out samples was becoming smaller with every time period, and, secondly, the majority of events occurred within the first 6 months.

For all the models, except for 'period-specific' ones ('Application score +ATS(s)', 'Application score + $Z(s)$'), the same set of parameters was applied to each of the hold-outs (H_1 to H_6). 'Period-specific' models contained sets of parameters, with each set corresponding to a certain time period.

For example, for Belgium the entry in column H_4 (4688) for 'Application score + Z(s)' model corresponds to

$$h^s(t) = e^{\beta'(s)x + \gamma'(s)z_{t-1}} h_0^s(t),$$

where $s = 4$. The parameter estimates for period 4 are used to predict $T \geq 5$ using the data available at period 4 (application score, ATS, delinquency status, percent repaid). The hold-out sample at period 4, H_4 , contained 4688 observations.

The results in Table 6.8 indicate that at the period of the first purchase the best prediction for $T \geq 2$ is obtained by including the Amount to Spend available right after the first purchase. This may be helpful in identifying customers that are going to make the second purchase immediately in the next period. However, for the Netherlands and Germany the difference is not as pronounced as for Belgium, so one can argue that early purchases can be successfully identified from the personal and the first purchase information.

As time progresses, behavioural information enhances the predictive power. For example (Table 6.8a), for Belgium the difference in the area under the ROC curve between 'personal, purchase' model and 'personal, purchase, behavioural information Z_{t-1}' (0.643 – 0.641)= 0.002 for holdout 2, whereas for holdout 6 the difference is (0.648-0.586)=0.062. This reinforces the view that over time the behavioural information becomes more important. The model with the best result for each holdout is identified in bold. For the full risk set (H_1) the best prediction is achieved by using personal and the first purchase information and initial ATS. Subsequently, the model that uses all behavioural information becomes a leader. For Belgium and the Netherlands 'specific time period' models outperform 'total observation period' models, but the German dataset shows the contrary result.

Table 6.8a) Predictive performance of national models with different levels of information. AUROC.

Belgium

Model	Hold-out					
	H_1 (7752)	H_2 (5725)	H_3 (5120)	H_4 (4688)	H_5 (4405)	H_6 (4167)
Personal	0.553	0.548	0.547	0.546	0.543	0.537
Personal+purchase	0.689	0.641	0.620	0.604	0.594	0.586
Personal+purchase +ATS_1	**0.715**	0.652	0.626	0.608	0.596	0.586
Personal+purchase +ATS_{t-1}	0.710	0.662	0.638	0.621	0.610	0.600
Application score+ATS(s)	0.706	**0.670**	0.654	0.649	0.641	0.635
Personal+purchase +Z_{t-1}		0.643	0.638	**0.658**	0.647	0.648
Application score+Z(s)		0.669	**0.658**	0.654	**0.654**	**0.650**

The Netherlands

Model	Hold-out					
	H_1 (9441)	H_2 (6227)	H_3 (4777)	H_4 (3987)	H_5 (3770)	H_6 (3583)
Personal	0.571	0.557	0.546	0.523	0.517	0.513
Personal+purchase	0.812	0.771	0.702	0.621	0.611	0.602
Personal+purchase +ATS_1	**0.820**	0.776	0.709	0.629	0.620	0.611
Personal+purchase +ATS_{t-1}	0.815	0.783	0.709	0.621	0.611	0.603
Application score+ATS(s)	0.819	0.784	0.712	0.645	0.640	0.641
Personal+purchase +Z_{t-1}		0.778	0.700	0.633	0.632	0.626
Application score+Z(s)		**0.791**	**0.722**	**0.671**	**0.668**	**0.668**

Germany

Model	Hold-out					
	H_1 (22799)	H_2 (18150)	H_3 (15851)	H_4 (14174)	H_5 (13186)	H_6 (12327)
Personal	0.553	0.548	0.541	0.537	0.534	0.533
Personal+purchase	0.774	0.752	0.728	0.704	0.693	0.684
Personal+purchase +ATS_1	**0.780**	0.757	0.731	0.708	0.695	0.687
Personal+purchase +ATS_{t-1}	0.774	0.759	0.738	0.716	0.704	0.694
Application score+ATS(s)	0.768	0.750	0.730	0.711	0.699	0.692
Personal+purchase +Z_{t-1}		**0.766**	**0.751**	**0.746**	**0.739**	**0.738**
Application score+Z(s)		0.764	0.743	0.734	0.726	0.728

Table 6.8b) Predictive performance of national models with different levels of information. Error rate.

Belgium

Model	Hold-out					
	H_1 (7752)	H_2 (5725)	H_3 (5120)	H_4 (4688)	H_5 (4405)	H_6 (4167)
Personal	39.34%	44.62%	45.86%	46.16%	46.68%	46.98%
Personal+purchase	30.70%	37.72%	40.40%	42.54%	43.18%	43.58%
Personal+purchase +ATS_1	**28.10%**	37.14%	40.70%	42.36%	43.82%	43.68%
Personal+purchase +ATS_{t-1}	28.20%	36.64%	39.60%	41.64%	42.28%	42.82%
Application score+ATS(s)	29.46%	**36.62%**	37.50%	39.08%	39.32%	39.84%
Personal+purchase +Z_{t-1}		38.50%	39.54%	38.44%	**38.54%**	**38.02%**
Application score+Z(s)		36.78%	**37.42%**	**38.40%**	38.72%	38.58%

The Netherlands

Model	Hold-out					
	H_1 (9441)	H_2 (6227)	H_3 (4777)	H_4 (3987)	H_5 (3770)	H_6 (3583)
Personal	35.58%	44.36%	46.98%	45.90%	44.50%	42.26%
Personal+purchase	21.16%	27.84%	34.16%	38.68%	38.36%	37.52%
Personal+purchase +ATS_1	**20.76%**	27.82%	34.34%	38.52%	38.20%	37.52%
Personal+purchase +ATS_{t-1}	21.00%	**27.52%**	34.38%	38.82%	38.14%	39.52%
Application score+ATS(s)	20.82%	27.82%	34.20%	36.76%	35.96%	34.38%
Personal+purchase +Z_{t-1}		28.48%	34.66%	37.08%	36.50%	35.12%
Application score+Z(s)		**27.52%**	**33.12%**	**35.62%**	**34.70%**	**33.78%**

Germany

Model	Hold-out					
	H_1 (22799)	H_2 (18150)	H_3 (15851)	H_4 (14174)	H_5 (13186)	H_6 (12327)
Personal	42%	46.08%	46.92%	46.84%	45.90%	44.34%
Personal+purchase	29.02%	32.00%	33.90%	35.06%	35.06%	34.10%
Personal+purchase +ATS_1	**28.62%**	31.44%	33.48%	34.78%	34.60%	33.68%
Personal+purchase +ATS_{t-1}	28.98%	31.18%	33.18%	34.28%	34.08%	33.48%
Application score+ATS(s)	30.02%	32.42%	33.50%	33.88%	33.92%	33.08%
Personal+purchase +Z_{t-1}		**30.36%**	**31.66%**	**31.36%**	**31.24%**	**30.02%**
Application score+Z(s)		31.08%	32.80%	32.98%	32.68%	31.08%

6.6. Good versus Bad

The initial formulation ignored the impact of default, although it was accounted for when the behavioural information that included delinquency status was included into the analysis. The increase in predictive power of models with behavioural information that was observed over time suggested that delinquent accounts were less likely to make further purchases, as one would expect.

It would be useful to know whether there is a difference between those who miss two consecutive payments within the observation period (Bads) and those who do not (Goods). Figure 6.6 gives baseline survival curves for Goods and Bads for three countries, and confirms that Goods have higher chances of making the second purchase.

Whilst survival curves for Goods exhibit different behaviour across the countries, especially for the Netherlands, survival curves for Bads nearly overlap, thus indicating certain similarity (Figure 6.7).

The sample was split into non-defaulters, Good, and defaulters, Bad. The Bads were divided into Bads after 2nd purchase and Bads before 2nd purchase. This last group would be considered censored in the competing risks formulation, but in fact, these are the customers that were delinquent, but then recovered and made the second purchase, and they constitute nearly 10% of all Bads and over 15% of Bads that make the second purchase.

Table 6.9 Number of 'Bads before/after the second purchase'

	Belgium	The Netherlands	Germany
Bads 1p	1450	3448	4380
Bads after 2p	1212	3482	3963
Bads before 2p	422	659	566
Total	3084	7589	8909

Figure 6.6 Baseline SDF by country for Good/Bad with 95% confidence intervals, personal data

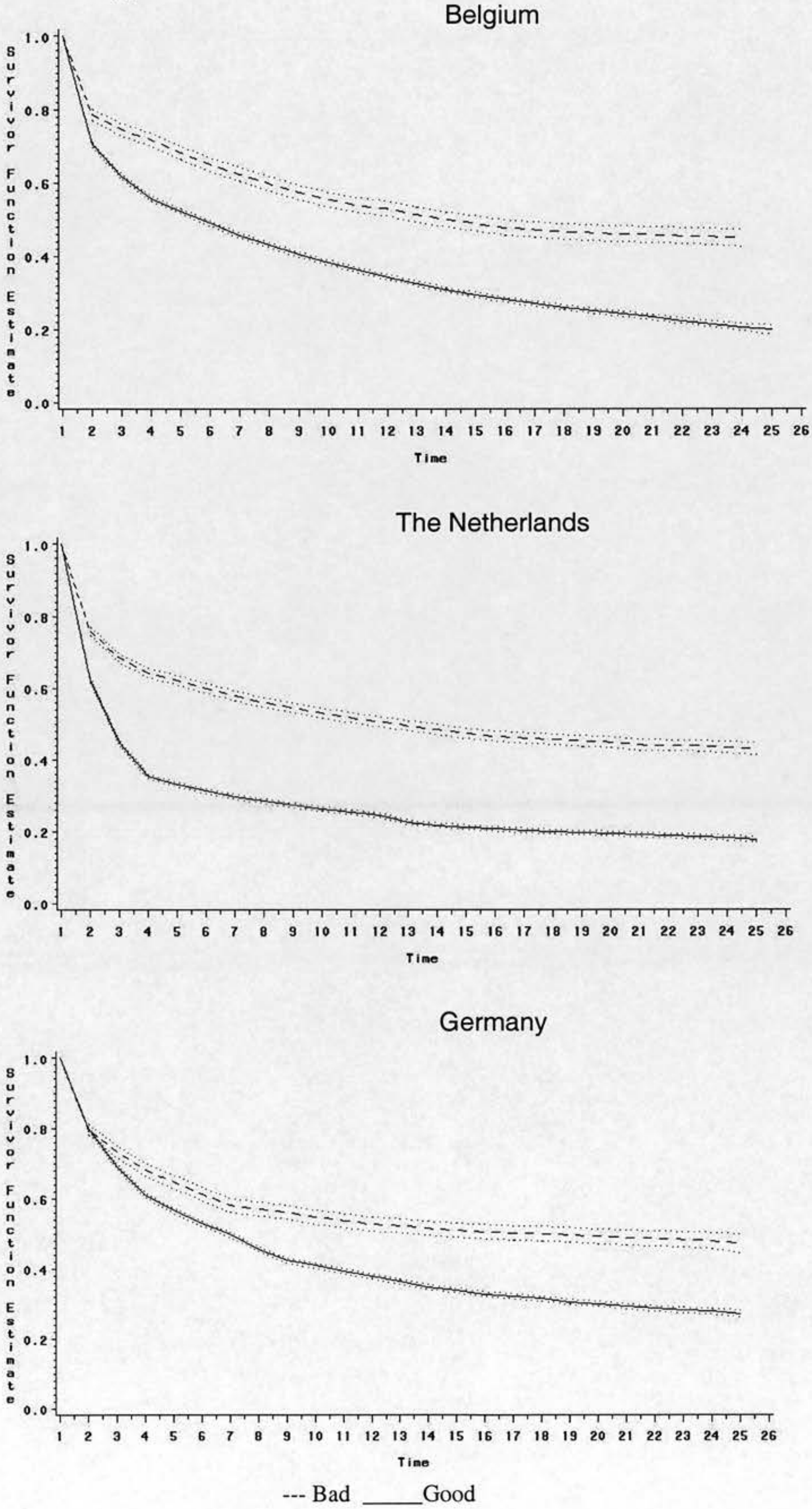
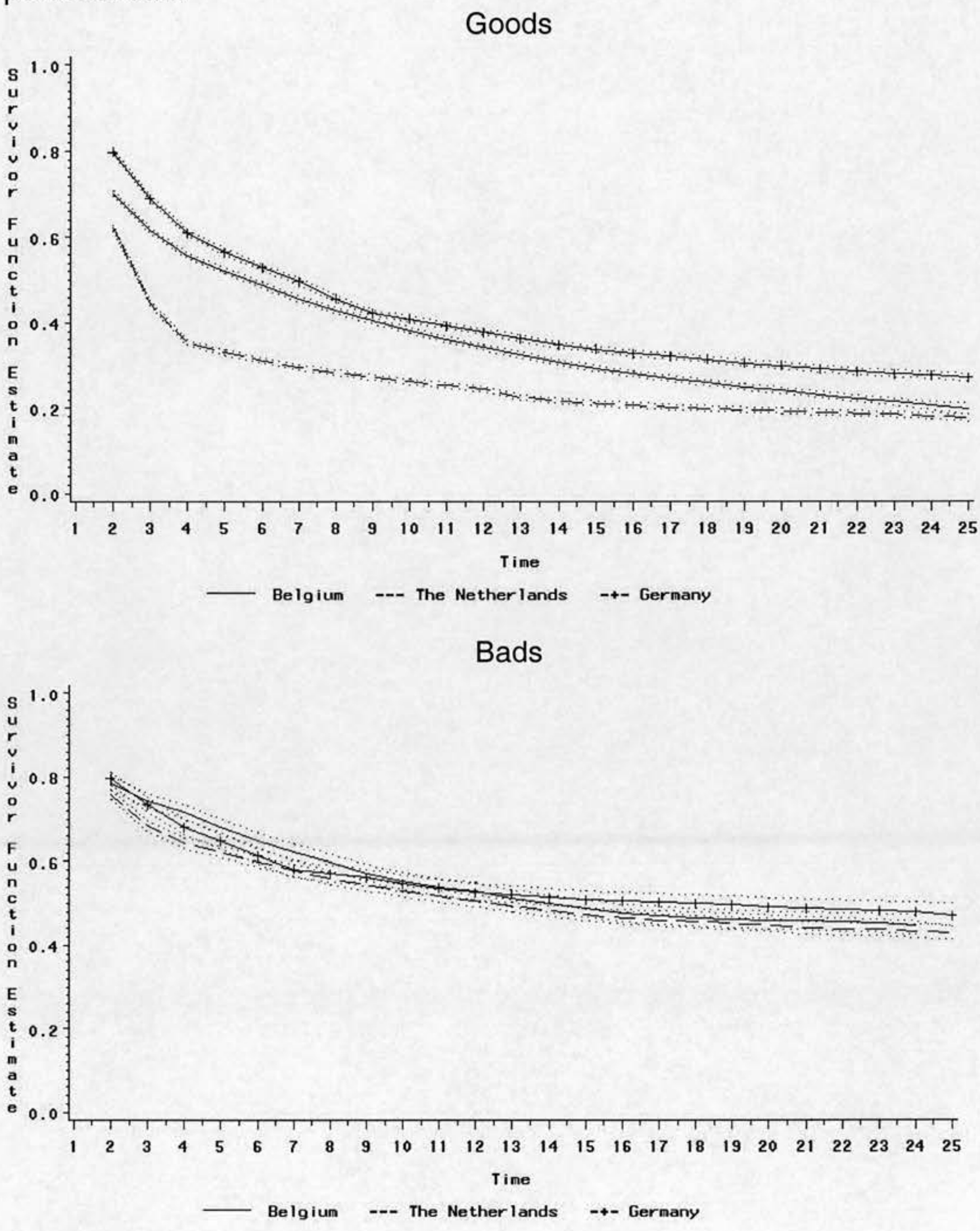


Figure 6.7 Baseline SDF for Good/Bad with 95% confidence intervals, personal data



Separate models were applied to each of the samples using personal, purchase and time-varying ATS variables. One should, however, treat the results with a degree of caution since as usual the number of defaulters is low. This will affect the quality of estimation and also the variables that will be deemed significant. Since the groups were small the estimation was done on training and hold-out samples combined.

The results are presented in Table 6.3 and Appendix A25, columns 7 to 9. There are differences between the models arising from each sample. As might be expected the models for the Goods are similar to that previously seen. In Belgium, both categories of Bads are associated with 'Phone given', whilst those that default before purchase have a new entry - number of children. Residential status, Industrial sector, Card and Credit insurance do not appear in the models for the Bad sample.

In the Netherlands phone indicator appears for 'Bads before'. Spouse's Age is important for both groups of Bads. No credit insurance signals higher chances of the next purchase for 'Bads after', whereas 'credit insurance 8'[13] indicates shorter time to the second purchase for all Bads. 'Bads after' that enter into the budget agreement with payment date other than the first day of the month, are faster buyers, and that is different from Goods and 'Bads before' falling into the same category.

In Germany, whether 'Phone given' is irrelevant for Goods and 'Bads before', but for 'Bads after', it indicates higher chances of the purchase. The full-time occupation is significant for 'Bads after'. 'Bads before' have higher chances of purchase, if they have a job in education, healthcare and their spouse is between 40 and 62, contrary to models estimated for the whole sample, where these categories were associated with lower chances of further purchase. The price of the first purchase shows a different trend for 'Bads before' - lower price corresponds to longer times to the next purchase.

So there does seem to be clear differentiation between the Good and the Bad payers, and equally there seem to be differences between the models for those who default before purchase and after. For credit control it means that longer estimated times to 2nd purchase signal not only a low purchase propensity, but also a higher default risk, so these accounts should be given extra attention. If missed payments are observed, the account may or may not recover. In the latter case it will be eventually written off. In the former case the lender will know that the customer will take longer time to the next purchase. So when (and if) the customer recovers, the time to the next purchase should be estimated using a different model. In case there was no 2nd purchase on the account, it should be a model for Bads who default before the 2nd purchase.

[13] No explanation for this category is available

6.7 Generic model performance

The generic model was developed on the aggregated dataset that was composed of the samples from the three countries. The composition of the generic sample is presented in Table 6.10. The proportions of different countries were kept roughly the same so none of the countries could dominate the generic model.

Table 6.10 Generic sample composition

	Good			**Bad**		
	2-nd purchase	**Censored**	**% Censored**	**2-nd purchase**	**Censored**	**% Censored**
Belgium	16640	6057	26.69%	1632	1449	47.03%
The Netherlands	15818	4060	20.42%	3371	2805	45.42%
Germany	15621	7475	32.36%	1546	1503	49.29%
Sub-total	48079	17592	26.79%	6549	5757	46.78%
Total	77977					

As before, 70% of 77977 observations were used as a training sample, the remaining 30% were reserved for testing the predictive ability of models. Several generic models were developed, each utilising different levels of information, so that they can be matched against the national models.

The results are summarised in Table 6.11. 'Personal' model demonstrates some of the trends that were observed in the national model: people in rented accommodation and recently at address are likely to make the purchase faster than home owners and those with longer years at address. The self-employed category is significant as before, and indicates lower chances for the next purchase. Number of kids and type of business are also significant and have a major effect on the time of the second purchase.

At the same time, there are some differences compared with the patterns of national models. Indicators of both home phone and employer's phone appear. Home Phone indicates higher chances of the purchase whereas employer's phone signals the contrary. Age, Spouse's Age and Marital Status are not selected into the model, which is not surprising given the controversial behaviour of these characteristics across the countries. Spouse's Age and Marital Status will appear at later stages when more information will be added and so will Time on Job. Only one age category will appear and only in one model.

Table 6.11 Generic models on different levels of information, parameter estimates

Variable	Model							
	Per-sonal	Personal, Purchase	Personal, Purchase, ATS_1	Personal, Purchase, ATS_{t-1}	From 2^{nd} period-more info	Personal, Purchase, ATS_{t-1} – Goods	Personal, Purchase, ATS_{t-1} -Bads after 2^{nd} p	Personal, Purchase, ATS_{t-1} – Bads before 2 p
	1	*2*	*3*	*4*	*6*	*7*	*8*	*9*
Employer's phone given	-0.062	-0.065	-0.082	-0.081		-0.086	-0.14	
Phone given	0.054	0.101	0.084	0.074	0.095		0.27	0.50
Kids: 0	-0.089	-0.140	-0.135	-0.142	-0.202	-0.171	-0.16	0.34
Kids: 1-2	-0.099	-0.155	-0.149	-0.154	-0.215	-0.190		0.42
Marital: Married								0.22
Marital: Single, divorced		-0.064	-0.058	-0.047	-0.057			
Res: Owner	-0.059				-0.101			
Res: Rented house	0.038	0.078	0.089	0.089	0.074	0.105		
Occup: Full-time	-0.063	-0.090	-0.090	-0.062	-0.075	-0.071		
Occup: Self-emp	-0.262	-0.227	-0.223	-0.199	-0.174	-0.173	-0.27	
Occup: Retired, housewife	-0.089	-0.145	-0.145	-0.160	-0.145	-0.151		
Type of business: Building	-0.208	-0.041	-0.064	-0.137	-0.050	-0.049		-0.31
Type of business: Healthcare	-0.133			-0.096	-0.110		-0.11	
Type of business: Industry	-0.130			-0.081	-0.074			
Type of business: Shop	-0.055			-0.053				
Age: under 22		-0.055						
Spouse age: no spouse		-0.045	-0.075	-0.069		-0.042		
Spouse age: under 26							0.17	
Time address: up to 6m	0.046							
Time address: 17y+	-0.067		-0.042	-0.044	-0.055	-0.060		
Time on job: 4-14y						-0.044		
Time on job: 14+		-0.043	-0.040	-0.044	-0.067	-0.082		
Time on job: Allowance								0.37

Table 6.11 Generic models on different levels of information, parameter estimates (continued from previous page)

	1	2	*3*	*4*	*6*	*7*	*8*	*9*
No card insurance		-0.258	-0.222	-0.254	-0.164	-0.206	-0.42	
Card insurance		-0.167	-0.109	-0.140	-0.107	-0.097	-0.26	
No credit insurance		0.058		0.047	0.073			-0.29
Product: heater, TV		-0.065	-0.073	-0.085	-0.087	-0.042	-0.08	
Product: hifi radio						0.058		
Product: kitchen items						0.048		
Product: card app, video		0.068	0.060	0.072	0.072	0.141		
Price: 0		0.386	0.278	0.327	0.391	0.372	0.21	0.32
Price: 1-180 EUR		0.497	0.366	0.343	0.399	0.379	0.32	0.41
Price: 181-450 EUR		0.399	0.254	0.282	0.353	0.330	0.13	-0.16
Price: 451-650 EUR		0.304	0.190	0.221	0.247	0.248	0.12	
Price: 651-800 EUR		0.218	0.122	0.154	0.190	0.180		
Price: 801-1100 EUR		0.127	0.086	0.098	0.121	0.120		
Pay date: 01		0.846	0.846	0.887	1.075	0.924	0.69	0.45
Pay date: 08		0.276	0.272	0.302	0.234	0.300	0.27	
Agreem: budget phone prop		0.590	0.607	0.500	0.500	0.398	0.72	
Agreem: budget other		0.446	0.530	0.428	0.420	0.275	0.68	
Agreem: deferred		0.480	0.449	0.355	0.290	0.234	0.72	1.36
ATS: over credit limit			-0.559	-1.211	-0.911	-0.780	-1.49	-2.33
ATS: 0-150 EUR			-0.168	-0.182	-0.171	-0.068	-0.30	-1.31
ATS: 151-500 EUR				-0.031				
ATS: 501-800 EUR			0.048		0.035			
ATS: 801-1200 EUR							0.12	
2+ missed payments					-7.119			
Percent repaid					0.741			

Table 6.12 Predictive performance of generic models with different levels of information. AUROC

Model	Hold-out											
	H_1 (23189)	H_2 (16769)	H_3 (14178)	H_4 (12510)	H_5 (11720)	H_6 (11009)	H_7 (10383)	H_8 (9790)	H_9 (9295)	H_{10} (8906)	H_{11} (8514)	H_{12} (7491)
Personal	0.567	0.545	0.524	0.508	0.508	0.508	0.510	0.518	0.520	0.516	0.512	0.503
Personal+purchase	0.726	0.704	0.671	0.637	0.629	0.623	0.616	0.604	0.592	0.592	0.592	0.586
Personal+purchase +ATS_1	0.734	0.708	0.674	0.642	0.635	0.625	0.619	0.606	0.593	0.593	0.592	0.588
Personal+purchase +ATS_{t-1}	0.721	0.698	0.663	0.628	0.619	0.613	0.606	0.594	0.581	0.580	0.579	0.573
Application score+ATS(s)	**0.735**	0.717	0.687	0.658	0.652	0.650	0.646	0.639	0.635	0.637	0.642	0.637
Personal+purchase +Z_{t-1}		0.715	0.696	**0.674**	**0.677**	**0.675**	**0.673**	0.667	0.660	0.663	0.663	0.660
Application score+Z(s)		**0.720**	**0.699**	**0.674**	0.674	0.673	0.671	**0.678**	**0.678**	**0.681**	**0.681**	**0.684**

Error rate

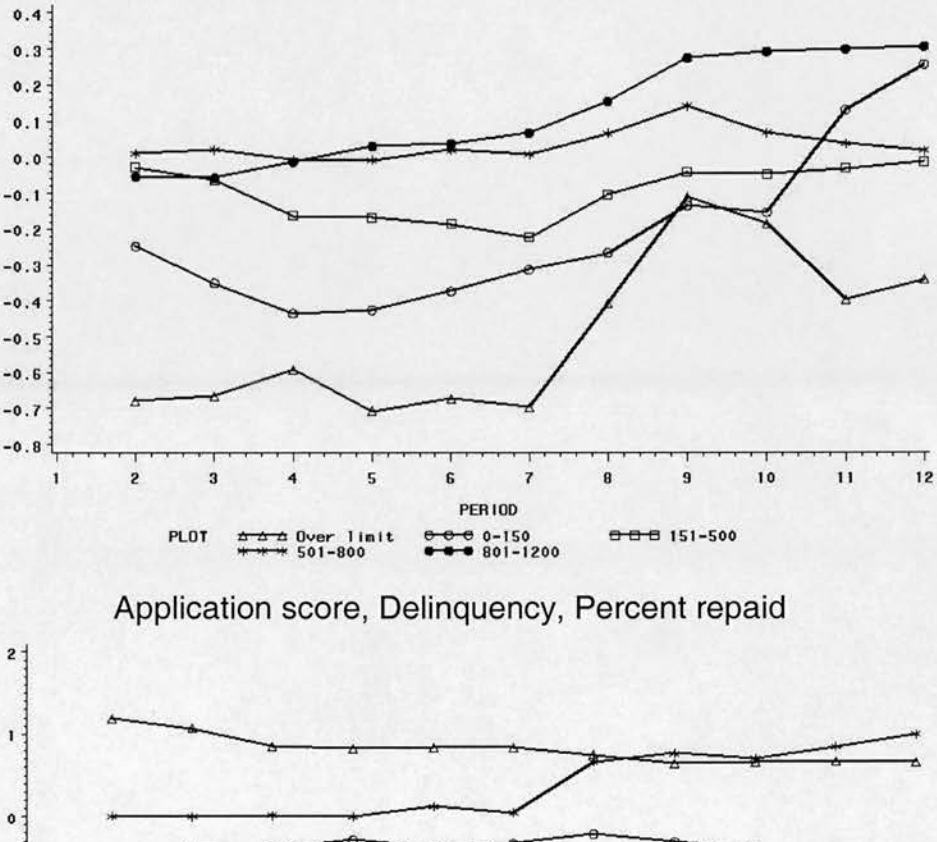
Model	Hold-out											
	H_1 (23189)	H_2 (16769)	H_3 (14178)	H_4 (12510)	H_5 (11720)	H_6 (11009)	H_7 (10383)	H_8 (9790)	H_9 (9295)	H_{10} (8906)	H_{11} (8514)	H_{12} (7491)
Personal	38.48%	45.58%	47.94%	49.04%	48.36%	46.60%	44.12%	40.72%	36.88%	33.80%	31.00%	29.18%
Personal+purchase	29.56%	33.70%	36.60%	38.68%	38.44%	37.38%	36.06%	34.92%	33.30%	30.92%	28.64%	27.48%
Personal+purchase +ATS_1	29.54%	33.74%	36.70%	38.36%	37.88%	36.96%	35.70%	34.38%	32.64%	30.40%	27.96%	26.54%
Personal+purchase +ATS_{t-1}	29.60%	34.14%	37.04%	38.98%	38.62%	37.68%	36.60%	35.34%	33.94%	31.48%	28.96%	27.74%
Application score+ATS(s)	**29.08%**	**33.56%**	36.44%	38.14%	37.28%	36.06%	34.90%	33.18%	31.32%	28.92%	26.50%	25.20%
Personal+purchase +Z_{t-1}		33.60%	**35.54%**	**36.62%**	**35.70%**	**34.60%**	**33.48%**	31.88%	30.34%	28.50%	26.22%	25.36%
Application score+Z(s)		33.66%	35.68%	37.38%	36.30%	35.02%	33.70%	**31.42%**	**29.46%**	**27.30%**	**25.12%**	**24.00%**

Purchase variables are highly significant as before, with the most important being Payment Date and Agreement Type. It is interesting to note that budget agreement type signals higher chances of further purchase than other agreement categories, obviously the result of Belgium and the Netherlands dominating over Germany.

ATS is also significant and shows the same trend, when higher value of credit availability corresponds to a quicker next purchase. Figure 6.8. shows that importance of application characteristics decreases with time, and importance of transactional data, on the contrary, increases, as was observed before.

Figure 6.8. Parameter estimates for behavioural data in generic model

ATS



Application score, Delinquency, Percent repaid

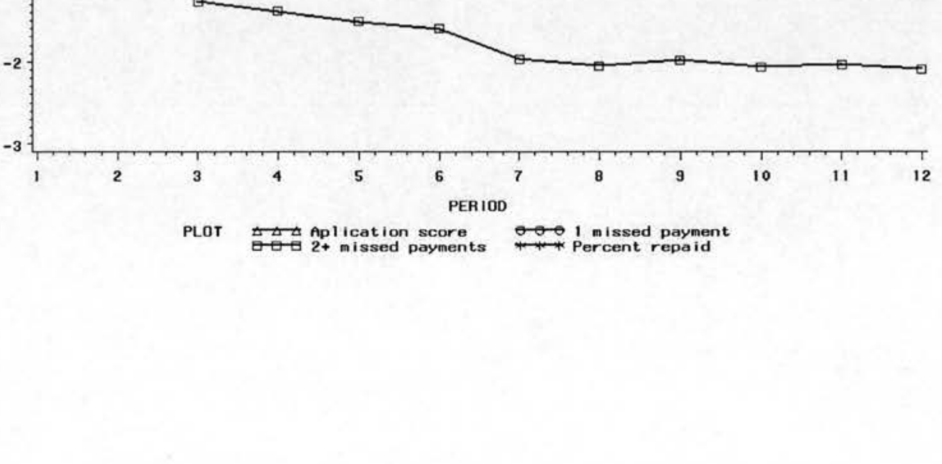


Table 6.13 shows the reduction in magnitude of Log Likelihood as more information is added into the analysis. Here the most significant reduction arises from incorporating the first purchase information in the analysis.

Table 6.13 Log Likelihood statistics for models with different levels of information

Model	Log L	df	Log L difference
No covariates	-402668		
Personal data	-402239	15	429
Personal+purchase	-398875	29	3364
Personal+purchase +ATS_1	-398542	31	333
Personal+purchase +ATS_{t-1}	-397916	35	625

The prediction was tested, first, on a set of generic aggregated hold-out samples, each sample corresponding to a relevant generic risk set for a time period T, using the same principle as in the previous section. Due to the larger risk sets, it was possible to extend the test to a longer period of time - 12 months (Table 6.12).

The results are in line with the prediction results demonstrated by the national models. Each level of information increases the predictive ability of models, with models using the complete range of application and behavioural data giving the best prediction (marked in bold). The importance of behavioural information increases with time, e,g, H_2 the difference between 'Personal + purchase' and 'Application score + Z(s)' model is 0.720-0.704=0.016, for H_{12} the same difference is 0.684-0.586=0.098. When behavioural information is limited to ATS, 'specific time period' models consistently outperform 'total observation period' models. With all behavioural information included, the superiority of the 'specific time period' approach becomes less evident, but even in the periods when 'total observation period' models give better prediction, the 'period-specific' models perform only slightly worse.

Each of the generic models was tested on the sets of national hold-out samples to see how their classification accuracy compares to that of national models across the countries. The results are shown in Figures 6.9-6.11. Each country demonstrates a distinctly different pattern.

Figure 6.9 Generic versus Belgian national model. Change in AUROC over time

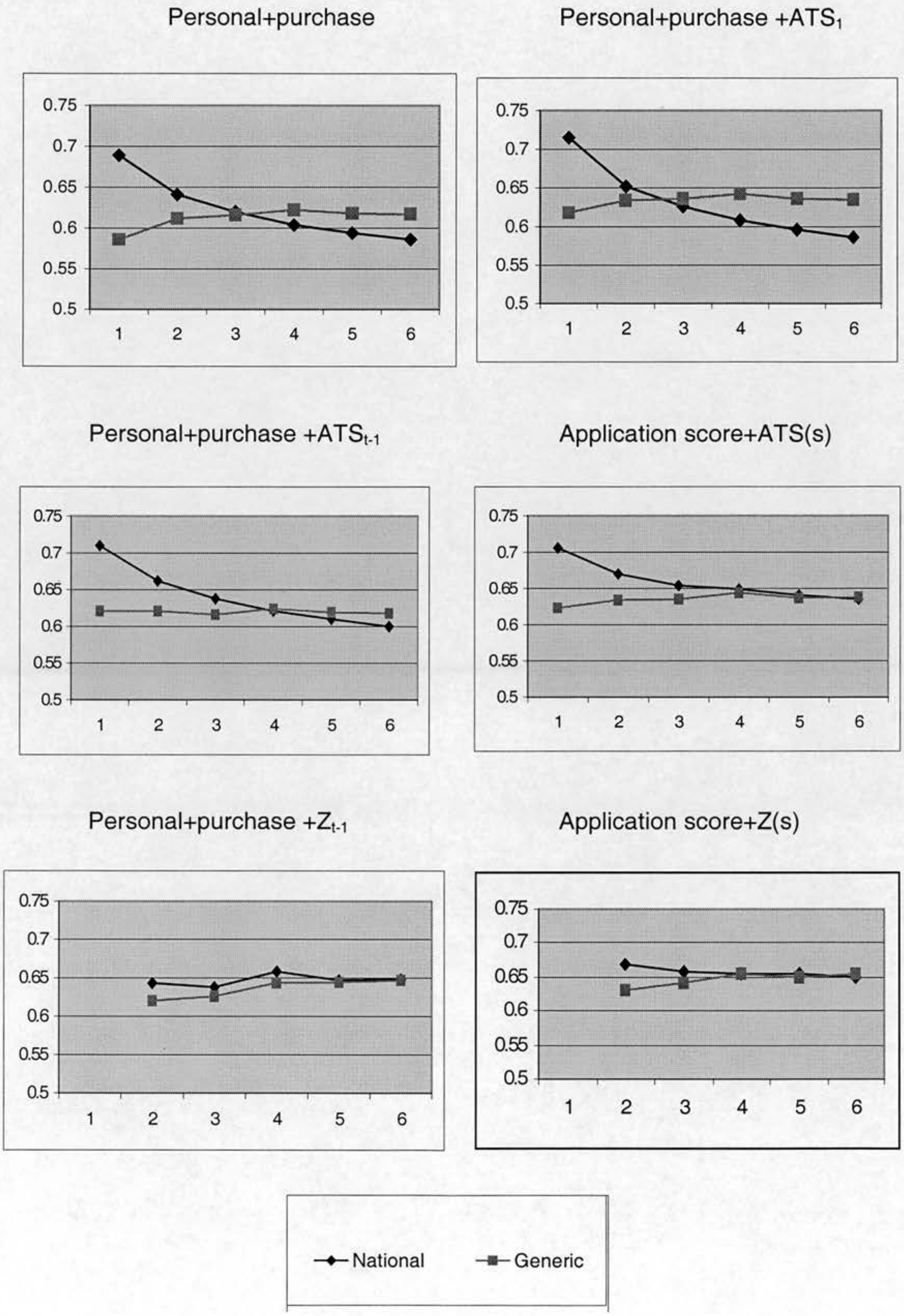


Figure 6.10 Generic versus Dutch national model. Change in AUROC over time

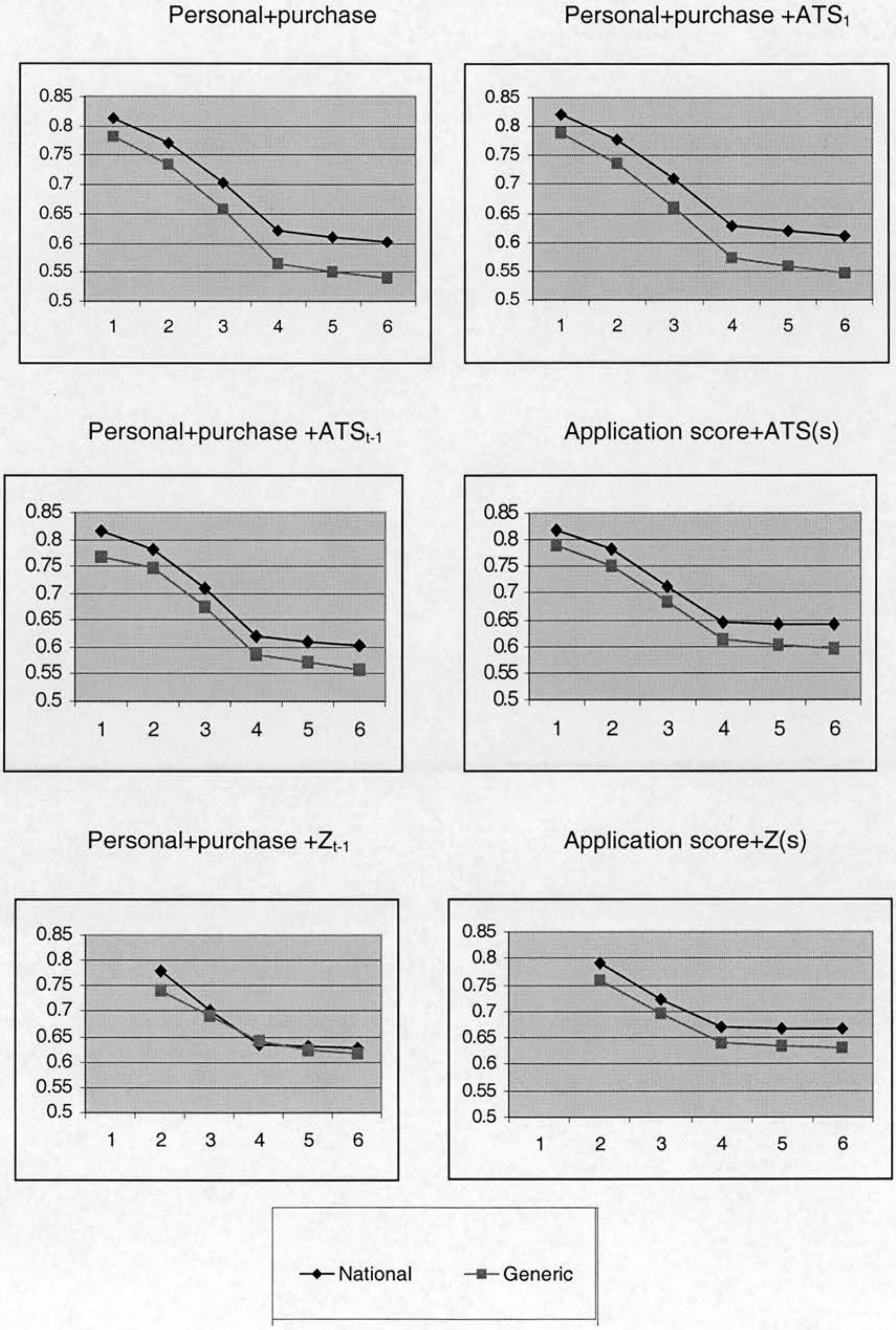
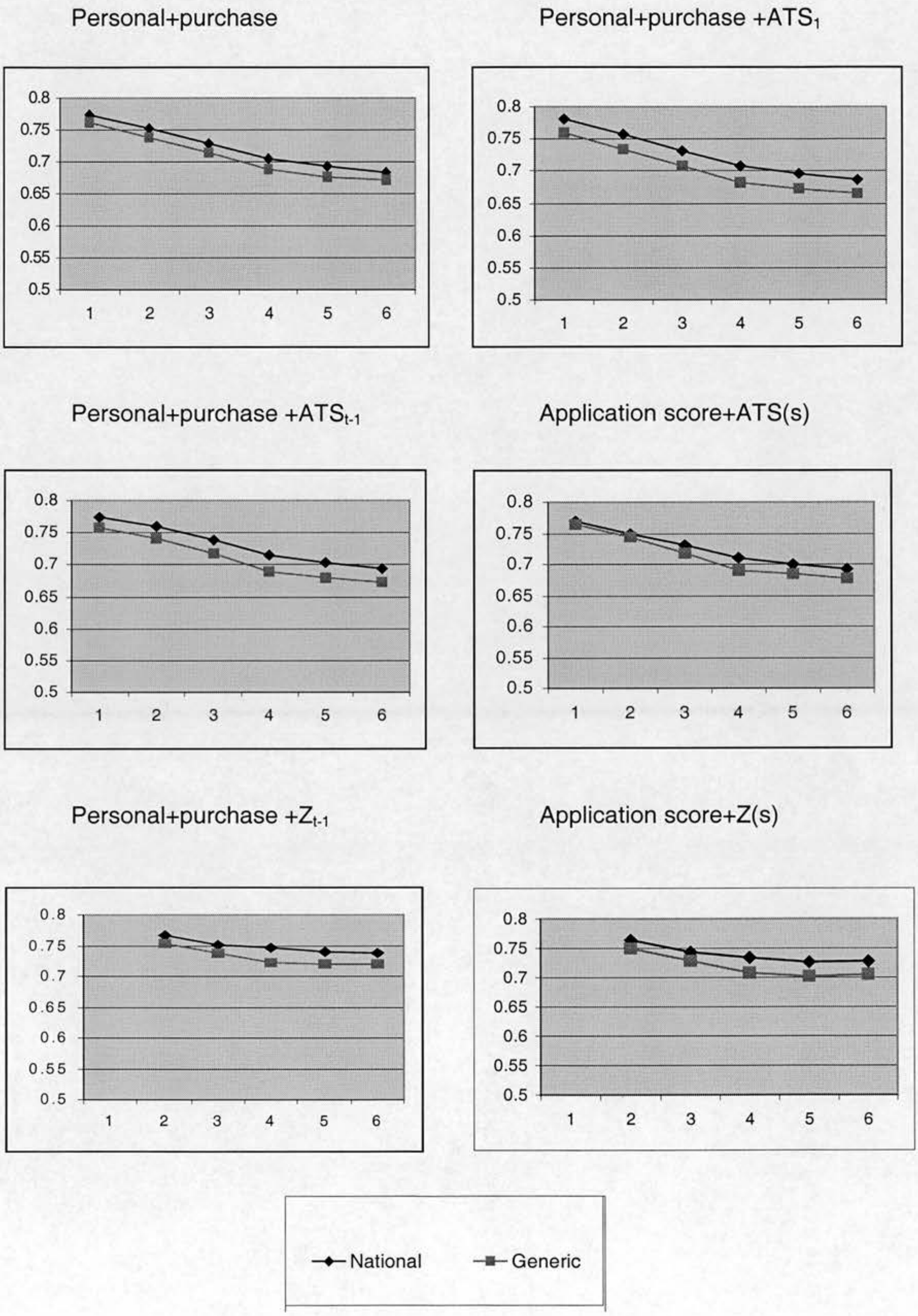


Figure 6.11 Generic versus German national model. Change in AUROC over time



For Belgium, the performance of national models, in general, deteriorates with time. In comparison, generic models show increasingly better classification accuracy, and at later time periods reach the level of performance of national models, or even outperform them. This can be attributed to the size of the samples on which the models were developed. The Belgian population was the smallest one in the analysis, so in this case the generic model had an advantage of a larger size of the training sample. Poor early performance of the generic model can be explained by the possible fact that early purchasers are different from the rest, and these differences are captured by the national model, but not by the generic model.

For the Netherlands, both generic and national models show considerable deterioration in performance. There is a sharp drop up to month 4, after this point the performance stabilises. The survivor function for the Netherlands in Figure 6.3 demonstrated exactly the same pattern: a rapid decline in the first 4 months, and then a smooth decrease after this point. The performance deteriorates in line with the changes in the size of the risk set over time. The national models consistently show superior performance, but the gap between the models becomes less as more information is added. The 'Personal+ purchase+ Z_{t-1}' demonstrate nearly identical performance.

Germany also demonstrates a decline in performance over time, but not as pronounced as the Netherlands or Belgium. The generic models give the closest performance to that of national models as compared to other two countries.

Overall, all three countries show that additional information makes the predictive performance more stable in time. It also diminishes the gap between the national and generic models, although this is not particularly evident for Germany. But in Germany the generic models perform reasonably well already at the level of 'personal+purchase' data. This emphasises the fact that generic scoring is highly dependent on the scope of information available. When all major predictors are included into the model, there will be a very slight difference between national and generic models, when this is not the case, national models will demonstrate a superior performance.

6.8 Conclusions

This Chapter presented a unique exploration of customer behaviour in using the retail card in three different countries. Modelling time to the second purchase when using a retail card was not addressed in the literature before. Knowledge of time to the second purchase will help in setting credit limits and retaining customers. The Chapter demonstrated for the first time how different types of data could be used and the value of each type in explaining the time to the next purchase.

Each level of information was found to enhance the modelling with a greater volume of variation accounted for and with better prediction results. It was found that purchase and transactional information was the most important, with variables that seem to have major effects being the Agreement Type and Amount to Spend. Amount to Spend was not explored by previous studies. It was established that ATS did not act as a first order Markov chain with the most recent value having the major effect. Whilst the customer required a positive credit balance before repeatedly using the card, it was also important how the customer got to this stage.

The importance of behavioural information increased with time and application information lost its discriminating power over time. Hence, the inclusion of behavioural information led to a better and stable performance over time.

The use of the card also indicated the likely behaviour of the individual. Defaulters were less likely to use the card than non-defaulters. This is partly imposed by lenders, since those in arrears with payments cannot use the card, hence they take longer time to the second purchase. There seemed to be differences between models built separately for Goods and Bads. For the Bads there was also a difference between those that defaulted before the second purchase and those that defaulted after the second purchase. Such differences may also help in identifying patterns so that action can be taken to protect the lender.

There were unique national patterns revealed across the countries with each level of information added. The performance of the generic model was not so impressive as when used to predict default as in Chapters 4 and 5. But it was shown that additional information diminished the gap in predictive ability between national and generic models. Hence, the need for further harmonisation of the information across the European countries is emphasised once again.

Chapter 7. Conclusions and extensions

This thesis has investigated the problem of using a single generic model to score several different populations. The problem has been viewed in the context of the integration of the European Union into a single market in consumer credit, with the following questions being addressed:

- Why generic credit scoring models may be needed in an integrated Europe?
- Are there any legislative restrictions on information that can be used in generic scoring models?
- What is the impact of these restrictions for lenders and borrowers?
- Are generic models competitive with customised models?
- Can we incorporate a time dimension into generic scoring?
- Can generic scoring be used in other applications, apart from predicting default?

This thesis presented the most thorough investigation of these questions to date, and many of these questions have not been previously addressed in the literature. The following sections will summarise the answers.

7.1 Why generic credit scoring models may be needed in an integrated Europe?

The European Union originated from an idea of combining several small and medium sized countries into one power, or even superpower, in order to enhance economic and political competitiveness of the Union members. The creation of internal single market in financial services is viewed as an important part of this process. It is envisaged that a free flow of credit across the nation states will foster competition and therefore, will strengthen the financial institutions and bring prosperity to consumers.

It should be noted that at present the integrated market in consumer credit is not yet a reality. But there is a great political desire and persistent work in progress to make this happen.

Financial institutions within the Union have a right to extend credit to residents of all Member States of the Union, although they are not obliged to do so. Those institutions that will offer credit on an international basis will gain the arising market opportunity. Obviously, there are risks associated with this opportunity. Credit applicants from foreign countries may present an unknown risk or a risk that is different to the risk of home country applicants.

There is always a possibility to segment the population on the basis of nationality/ country of residence and to apply tailored risk assessment procedures to each segment. This may be an expensive and complicated strategy, with potentially fifteen different models to develop and maintain for existing EU members, plus more models for new members that will join the EU in the future. An alternative is to use a generic model, or to split the whole European market into three or four more or less homogeneous segments.

Apart from the common sense perspective, there may be a legal requirement to assess the credit risk without making distinctions between residents of different European countries. One cannot say definitively that there is such a requirement at present, since there is no explanation at the European level of what exactly constitutes discrimination in credit scoring. That is why the principle of non-discrimination on the grounds of nationality, the EU fundamental principle, is open to interpretation by the national authorities. And history shows that this principle can be interpreted as a prohibition to use nationality as a variable in credit scoring models. This brings us to the next question whether there are other variables that potentially cannot be included into credit scoring models.

7.2 Are there any legislative restrictions on information that can be used in generic scoring models?

Whilst in the USA there is a specific legal act that gives the list of characteristics that cannot be used when making lending decisions (ECOA), in Europe the situation is far from straightforward. Restrictions on information should be inferred from general anti-discrimination and data protection provisions that differ from country to country, as has been shown in Chapter 2 of this thesis. So lenders extending credit to applicants outside the lender's country of registration should take into account all these varying requirements.

In general, the law does not distinguish between 'subjective' and 'objective' discrimination, instead the distinction is made between 'purpose' and 'effect'. The use of certain information in credit scoring models can be regarded as an action with the 'purpose' of separating certain groups from the others. That is why 'non-discrimination' is often interpreted as equivalent to 'making no distinction', and therefore, the use of information that allows an organisation to make such distinction is prohibited.

Race, nationality, and gender are already subject to anti-discrimination regulations at the European Union level. Disability and religion are included into the scope of law in many countries. Age is most likely to follow. But there is significant scope for different interpretations, so case law becomes extremely important. We are not aware of any cases against lenders in Europe, apart from the case in France that was discussed in Chapter 2.

Whilst the question of what exactly can be used in credit scoring models in Europe remains open, the related question is whether one should worry about the prohibition of information at all. Is there any effect from prohibition?

7.3 What is the impact of restrictions on information for lenders and borrowers?

This thesis reviewed a substantial number of studies on the effects of prohibition on information in credit scoring models and concluded that prohibition was detrimental for both lenders and borrowers. If prohibited information has a significant association with the default probability, removing this information from the model will have the following implications. First, it has been shown that some protected groups are good credit risks, but without the relevant information in the model they are not given 'credit' for being good. Second, the acceptance rates are higher for protected groups if the lender has an ability to segment or separate these groups, especially if these groups are significantly different from the rest of population. Third, the lenders may not be able to achieve the same level of classification accuracy, which will increase the level of bad debt, and the resulting increased cost of credit will be passed on to the consumer.

7.4 Are generic models competitive with customised models?

Previously published research showed the superiority of customised models. This thesis used a unique dataset that was particularly suited for a comparison between customised and generic models. The generic model was found to perform well. In predicting default the generic model demonstrated only a very slight inferiority in prediction as compared to national models developed on the same set of characteristics.

Several reasons may serve as an explanation for this result. First, only one type of credit product was used in the analysis. Second, the investigated countries were close in geographical and socio-economic terms. General trends in behaviour of good/bad classes were traced across the countries, and the differences observed were compensated by the flat maximum effect. Third, the generic model was developed on a heterogeneous sample that equally represented the national populations. Such an approach is not without limitations. If one assumes that the distinction between countries is not allowed, then it will not be possible to construct a sample that will represent the countries equally. Still we believe that the approach taken in this thesis was justified, otherwise the smaller nations would be dominated by a larger one and there would be a 'German' model instead of a 'generic' one.

It should be noted that additional information increases the predictive ability of models, so the performance of generic models depends to a larger extent on the scope of information available. Therefore, harmonisation of data and data collection procedures becomes important.

7.5 Can we incorporate a time dimension into generic scoring?

It has been shown that there was practically no difference in classification accuracy achieved by logistic regression and survival analysis modelling techniques in predicting default. Further, different parametric AFT approaches and the semi-parametric Cox PH model produced very similar results. This was observed both for generic and national models. Survival analysis generic models predicted as well as the logistic regression models.

It was suggested that exponential distribution offered a suitable fit for the analysed dataset, therefore, implying that the hazard was constant throughout 25 months. That is why survival analysis and logistic regression demonstrated close results in prediction. This is different from the results of previous studies conducted on fixed term loans, suggesting that revolving credit may have a more random character.

Since both survival analysis and logistic regression demonstrate similar prediction, the decision as to which approach should be used depends on the additional advantages each approach can offer. The main advantage of survival analysis is its ability to generate predictions that give the probability of 'survival' in any specific time period, without the necessity of re-estimating the model to fit this time period, as is the case with logistic regression.

In spite of the observed differences in survival patterns between the countries, the survival analysis generic models performed well, therefore, a time dimension can be incorporated into generic scoring.

7.6 Can generic scoring be used in other applications, apart from predicting default?

The estimates of time to default can be used as a basis for estimating the profit a customer will generate, see Stepanova (2001). But in the context of revolving credit it is necessary to predict the repeated usage of the product in order to calculate the profit. It was shown that generic scoring can be used to model not only time to default, but also time to the second purchase that was selected as a measure of usage.

Several types of information (personal, purchase and behavioural data) were included into analysis sequentially and the value of each type was investigated. Purchase and transactional information was found to be most important with the value of transactional information increasing over time.

The differences in purchasing behaviour between delinquent and non-delinquent accounts were observed, and also between accounts that get into delinquency before the second purchase and after it. It is interesting to note that the survival curves for delinquent accounts appeared to be very similar across the three countries, whilst for non-delinquent accounts they demonstrated significant differences.

The generic models performed well, but the difference in predictive accuracy between them and national models was more pronounced than in predicting default. However, inclusion of additional information, behavioural information in particular, brought the performance of generic models closer to that of national models. This finding emphasises once again the importance of information. The availability of information is crucial and, hence restrictions on its use are anti-consumer, and harmonisation of information across Europe is highly desirable.

7.7 Implications of the research

7.7.1 Implications for academics

Following the paper by Platts and Howe (1997), this thesis opens a new direction of inquiry which can be defined as comparative or cross-cultural credit scoring. In comparison to the above mentioned paper, the thesis presents a thorough investigation of the national credit risk patterns in three European countries. A unique dataset has been used in the analysis with a tighter control over the type of credit product than in the previous study, and the combination of countries explored by the thesis has not been investigated before.

The thesis provides the first comparison of legal restrictions on the data used in credit scoring between the USA and EU and within the EU, and investigated the impact of these restrictions for lenders and borrowers. It also gives a thorough review of literature on generic models and related issues.

In contrast to previous findings the generic model has demonstrated only a slight loss in predictive accuracy compared to national models suggesting that generic scoring will work successfully for certain groups of countries provided there is enough information in common.

The generic model was tested under different modelling techniques and in different applications. The thesis investigated time to default and time to the second purchase across the three countries. This type of comparison has not been reported in the published literature. Exponential distribution was found to be suitable for modelling time to default, implying that revolving credit may have a more random character as compared to fixed-term loans. But this finding needs to be tested on other datasets.

Time to the second purchase has not been investigated by previous research in the context of retail credit. The thesis has examined the value of different types of information in modelling time to the second purchase and how this value changes over time. It also found differences in purchasing behaviour between delinquent and non-delinquent accounts, and that the majority of defaults and purchases happened relatively early in the credit lifecycle. These findings can serve as a basis for further research of revolving credit.

7.7.2 Implications for business community

The thesis offers a valuable source of information for lenders that extend credit across the EU or are going to do so. To begin with, there are considerations of legal compliance of credit scoring models. Lenders offering credit to borrowers in other EU countries have to ensure that they comply with regulations of these MS.

At present the use of generic models is limited to situations when there is no data available for building a customised model (e.g., new product development, expansion into a new market). This thesis has demonstrated that for certain countries national models can be replaced with a generic model with only minor loss in predictive accuracy. The loss in predictive power depends on the scope of information included into the model, therefore, lenders should consider harmonising application data across the countries that they offer credit to.

It also has been shown that switching from logistic regression to survival analysis leads to no loss in predictive power of national and generic models, but at the same time offers a possibility for profit scoring. It also allows for debt provisioning and cash flow forecasting. It provides insight into timing of events and thus gives information when certain actions should be planned.

With the growing competition among lenders, retention of customers becomes increasingly important. This thesis demonstrated that 'purchase scoring' can be helpful in identifying 'slow spenders' and in providing information for setting and adjusting credit limits. Besides, the observed differences between delinquent and non-delinquent account, as demonstrated in Section 6.6, suggest that if a customer falls behind with payments but then recovers, his/her subsequent behaviour should be assessed by a different model.

Overall, it was shown that generic scoring is a viable option suitable for different applications.

7.7.3 Implications for policy-makers

This thesis identified some important aspects that have to be addressed to allow for effective risk assessment of applicants for credit from different EU countries. Throughout the thesis the need for further harmonisation of information used in credit scoring has been constantly emphasised. In particular, harmonisation of CRA data should be considered. If lenders are unable to receive and match the CRA records from different countries, this can create a serious obstacle to a free flow of credit across nation states.

Equally important is the need to reconsider what constitutes discrimination in credit scoring, and what is an appropriate balance between 'equal treatment' and 'responsible lending'. The absence of coherent and clear explanation of what can and cannot be used in credit scoring models does not contribute to creation of an integrated market either.

But a warning should be given against the use of a straightforward prohibitive approach in eliminating discrimination in credit markets, since it was shown that bans on information in credit scoring proved to be detrimental to both lenders and borrowers.

7.8 Further research

There are broadly four directions that the research conducted in this thesis can be extended. First, there is a need to explore more European countries in order to identify similarities and differences in credit risk patterns across Europe.

Second, it was shown that both types of events investigated (default and the second purchase) occur early in the lifetime of an account. One of the possible approaches to modelling early events would be to use frailty models which assume that there is some unobserved quality that distinguishes these events from the rest. A random effect variable is introduced that describes excess risk for certain groups of borrowers. The penalised Cox model can be used as a modelling tool (Therneau and Grambsch (2000)). This can be an alternative approach to the parametric one presented in Hand and Vinciotti (2003).

Third, it would of interest to further investigate the propensity of credit applicants to use the revolving credit product. This thesis looked into one possible approach, whilst there are other approaches possible: estimating the number of purchases, the total value of purchases, modelling multiple events. It is also necessary to explore the ways to combine modelling of default and purchase. Multi-state modelling (Therneau and Grambsch (2000)) can be a possible way forward in this direction.

Fourth, the performance of generic models is sensitive to the scope of information available, hence, exploration of additional sources of predictors can be pursued. One obvious source of information is CRA, and the research into how the CRA information can be harmonised across the EU is highly important. The absence of such information creates a major constraint on the way towards the integrated market in consumer credit. Another possible way to incorporate more information is to include economic indicators into generic models and to investigate their effect on performance.

Bibliography

1. Allison, P.D. (1995) *Survival analysis: using the SAS system,* SAS Publishing, Cary, NC, USA.

2. Allison, P.D. (1999) *Logistic Regression Using the SAS System: Theory and Applications*, SAS Institute, Inc.

3. Altman, E.I., Avery, R.B., Eisenbeis, R.A., and Sinkey, J.F., Jr. (1981) *Applications of Classification Techniques in Business, Banking, and Finance*, Greenwich, CT: JAI Press.

4. Anderson, P.K. and Gill, R.D. (1982) Cox's regression model for counting processes: a large sample study, *Annals of Statistics*, 10, 1100 - 1120.

5. Andreeva, G., Ansell, J.I. and Crook, J.N. (2004) Impact of anti-discrimination laws on credit scoring, *accepted for publication in Journal of Financial Services Marketing*.

6. Ansell, J.I. and Phillips, M.J. (1994) *Practical Methods for Reliability Data Analysis*, Clarendon Press, Oxford.

7. Ansell, J.I., Archibald, T., and Harrison, T. (2001) Lifestage, lifestyle and the impact on the client propensity in financial services, *The University of Edinburgh Credit Research Centre Working Paper Series*, 01/3, Edinburgh.

8. Aster, R. and Seidman, R. (1996) *Professional SAS programming secrets*, McGraw-Hill.

9. Avery, R.B. (1981) Indirect Screening and the Equal Credit Opportunity Act, *Research Papers in Banking and Financial Economics*, Washington: Board of Governors of the Federal Reserve System, Division of Research and Statistics, Financial Studies Section.

10. Avery, R.B. (1982) Discrimination in Consumer Credit Markets, *Research Papers in Banking and Financial Economics*, Washington: Board of Governors of the Federal Reserve System, Division of Research and Statistics, Financial Studies Section.

11. Avery, R.B., Bostic, R.W., Calem, P.S., and Canner, G.B. (2000) Credit scoring: statistical issues and evidence from credit-bureau files, *Real Estate Economics*, 28, 3, 523 - 547.

12. Baesens, B. (2003) Developing intelligent systems for credit scoring using machine learning techniques, PhD Thesis, Katholieke Universiteit, Leuven.

13. Baker, S. and Kramer, B. (2001) Good for women, good for men, bad for people: Simpson's paradox and the importance of sex-specific analysis in observational studies, *Journal of Women's Health & Gender-Based Medicine*, 10, 9, 867 - 872.

14. Bamber, D. (1975) The area above the ordinal dominance graph and the area below the receiver operating characteristic graph, *Journal of Mathematical Psychology*, 12, 387 - 415.

15. Banasik, J.L., Crook, J.N., and Thomas, L.C. (1996) Does scoring subpopulations make a difference?, *International Review of Retail, Distribution and Consumer Research*, 6, 2, 180 - 195.

16. Banasik, J., Crook, J.N., and Thomas, L.C. (1999) Not if but when will borrowers default, *Journal of Operational Research Society*, 50, 1185 - 1190.

17. Banasik, J.L., Crook, J.N., and Thomas, L.C. (2001a) Sample selection bias in credit scoring models, Working Paper Series, 01/5, Credit Research Centre University of Edinburgh.

18. Banasik, J.L., Crook, J.N., and Thomas, L.C. (2001b) Scoring by usage, *Journal of the Operational Research Society* , 52, 997 - 1006.

19. Banasik, J.L. and Crook, J.N. (2002) Does reject inference really improve the performance of application scoring models?, Working Paper Series, 02/3, Credit Research Centre University of Edinburgh.

20. Banasik, J.L. and Crook, J.N. (2003) Reject inference and lean models, in *Credit Scoring and Credit Control VIII Conference Proceedings*, Credit Research Centre, The University of Edinburgh.

21. Banton, M. (1994) *Discrimination*, Open University Press.

22. Barron, J.M., Ellienhausen, G., and Staten, M.E. (2000) Monitoring the household sector with aggregate credit bureau data, *Business Economics*, 35, 1, 63 - 77.

23. Becker, G.S. (1971) *The Economics of Discrimination*, University of Chicago Press, Chicago.

24. Belsley, D.A., Kuh, E., and Welsch, R.E. (1980) *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, John Wiley & Sons.

25. Bennet, C.J. (1992) *Regulating privacy: data protection and public policy in Europe and the United States*, Cornell University Press.

26. Benston, G.J. (1997) Discrimination in Financial Services: What Do We Not Know? *Journal of Financial Services Research*, 11, 1, 209-213.

27. Betten, L. and Grief, N. (1998) *EU law and human rights*, Harlow : Longman.

28. Bostic, R.W. (1997) The role of race in mortgage lending: revisiting the Boston Fed Study, Finance and Economics Discussion Paper Series, 02, Board of Governors of the Federal Reserve System.

29. Boyle, M., Crook, J.N., Hamilton, R., and Thomas, L.C. (1992) Methods for credit scoring applied to slow payers, in *Credit Scoring and Credit Control*, ed. Thomas, L. C., Crook, J. N., and Edelman, D. B., Oxford University Press, Oxford.

30. Boyes, W.J., Hoffman, D., and Low, S. (1978) Lender Reactions to Information Restrictions: The Case of Banks and ECOA, *Journal of Money, Credit and Banking*, 18, 211 - 219.

31. Breslow, N.E. (1974) Covariance analysis of censored survival data, *Biometrics*, 30, 89 - 99.

32. Brimelow, P. and Spencer, L. (1993) The Hidden Clue, *Forbes*, January 4, 1993, p.48.

33. Carow, K.A. and Staten, M.E. (1999) Debit, credit or cash: survey evidence on gasoline purchases, *Journal of Economics and Business*, 51, 409 - 421.

34. CBS (2003) The Year in Figures 2003, Statistics Netherlands, http://www.cbs.nl/en/publications/articles/general/yif/yif2003.pdf.

35. Chandler, G.G. and Ewert, D.C. (1976) Discrimination on the Basis of Gender Under the Equal Credit Opportunity Act, Working Paper, 8, West Lafayette, IN, Purdue University, Credit Research Center.

36. Chandler, G.G. and Johnson R.W. (1992) The benefit to consumers from generic scoring models based on credit reports, *IMA Journal of Mathematics Applied in Business & Industry*, 4, 61 - 72.

37. Chandler, G.G. (1998) Generic and customized scoring models: a comparison, in *Credit Risk Modeling*, ed. Mays, E., Glenlake Publishing Company, LTD.; Fitzroy Dearborn Publishers, Chicago.

38. CNIL (1998) Délibération n° 98-101 du 22 décembre 1998 portant modification de la recommandation relative à la gestion des crédits ou des prêts consentis à des personnes physiques par les établissements de crédit, http://www.cnil.fr/publi/index.htm.

39. COFIDIS (2003) Le codran de COFIDIS, 18, http://www.cofidis.com/us/us_cad/pdf/cadran_18.pdf.

40. Collett, D. (1994) *Modelling Survival Data in Medical Research*, Chapman & Hall.

41. Council of the EU (1987) Council Directive 87/102/EEC of 22 December 1986 for the approximation of the laws, regulations and administrative provisions of the Member States concerning consumer credit, *Official Journal of the European Communities*, L 042, 0048 - 0053.

42. Council of the EU (1989) Second Council Directive 89/646/EEC of 15 December 1989 on the coordination of laws, regulations and administrative provisions relating to the taking up and pursuit of the business of credit institutions and amending Directive 77/780/EEC, *Official Journal of the European Communities*, L 386, 0001 - 0013.

43. Council of the EU (1990) Council Directive 90/88/EEC of 22 February 1990 amending Directive 87/102/EEC for the approximation of the laws, regulations and administrative provisions of the Member States concerning consumer credit, *Official Journal of the European Communities*, L 061, 0014 - 0018.

44. Council of the EU (2000a) Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin, *Official Journal of the European Communities*, L180, 0022 - 0026, http://europa.eu.int/eur-lex/pri/en/oj/dat/2000/l_180/l_18020000719en00220026.pdf.

45. Council of the EU (2000b) Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation, *Official Journal of the European Communities*, L 303, 0016.

46. Council of the EU (2001) Council Resolution of 26 November 2001 on consumer credit and indebtedness, *Official Journal of the European Communities*, C 364, 0001 - 0001, http://europa.eu.int/eur-lex/pri/en/oj/dat/2001/c_364/c_36420011220en00010001.pdf.

47. Cox, D.R. (1972) Regression models and life-tables (with discussion), *Journal of Royal Statistical Society, Series B*, 74, 187 - 220.

48. CRA (1977) The Community Reinvestment Act, *12 U S Code 2901*.

49. Crook, J.N., Hamilton, R., and Thomas, L.C. (1992) Credit card holders: users and nonusers, *Service Industry Journal*, 12 , 251 - 262.

50. Crook, J.N. and Thomas, L.C. (1992) The degradation of the scorecard over the business cycle, *IMA Journal of Mathematics Applied in Business & Industry*, 4, 111- 123.

51. Crook, J.N. (1997) Application credit scoring: An overview, *Journal of Financial Services Marketing*, 2, 2.

52. Crook, J.N. (1999) Who is discouraged from applying for credit, *Economics Letters*, 65, 2, 165 - 173.

53. Crowder, M. (2001) *Classical Competing Risks*, Chapman & Hall/CRC.

54. Data Protection (2000a) Data Protection Commissioner & lending credit industry move closer together on third party credit data (Press Statement), http://wood.ccta.gov.uk/dpr/dpdoc1.nsf.

55. Data Protection (2000b) Data Protection Commissioner welcomes credit industry proposals on way forward for third party credit data (Press Statement), http://www.dataprotection.gov.uk/dpr/dpdoc1.nsf.

56. Davis, R.H., Edelman, D.B., and Gammerman, A.J. (1992) Machine-learning algorithms for credit-card applications, *IMA Journal Of Mathematics Applied in Business & Industry*, 4, 61 - 72.

57. Dawes, R.M. (1979) The robust beauty of improper linear models in decision making, *American Psychology*, 34, 7, 571 - 582.

58. Day, T.E. and Liebowitz, S.J. (1998) Mortgage Lending to Minorities: Where's the Bias?, *Economic Inquiry*, 36, 1, 3 - 28.

59. DDA (1995) Disability Discrimination Act, http://www.hmso.gov.uk/acts/acts1995/1995050.htm.

60. DeLong, E.R., DeLong, D.M., and Clarke-Pearson, D.L. (1998) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach, *Biometrics*, 44, September 1988, 837 - 845.

61. De Nederlandsche Bank (2003) T2.2.1 Loans from MFIs to the private sector, breakdown by sector, original term and instrument, http://www.statistics.dnb.nl/indexuk.html.

62. Desai, V.S., Conway, D.G., Crook, J.N., and Overstreet, G.A.J. (1997) Credit-scoring models in the credit-union environment using neural networks and genetic algorithms, *IMA Journal of Mathematics Applied in Business & Industry*, 8, 323 - 346.

63. Draft Treaty (2003) Draft Treaty establishing a Constitution for Europe, *Official Journal of the European Communities*, C 169.

64. Durand, D. (1941) Risk Elements in Consumer Instalment Financing, Studies in Consumer Instalment Financing, 08, New York: National Bureau of Economic Research.

65. ECOA (1976) Regulation B - Equal Credit Opportunity, *15 U S Code*, §§ 1691-1691f, http://www.ftc.gov/ogc/stat3.htm.

66. Efron, B. (1974) The efficiency of Cox's likelihood funtion for censored data, *Journal of American Statistical Association*, 76, 312 - 319.

67. Elliehausen, G.E. and Durkin, T.A. (1989) Theory and Evidence of the Impact of Equal Credit Opportunity: an Agnostic Review of the Literature, Monograph, 28, Krannert Graduate School of Management, Purdue University.

68. European Commission (1995) Report on the operation of Directive 87/102/EEC for the approximation of the laws, regulations and administrative provisions of the Member States concerning consumer credit, COM(95) 117 final, Brussels.

69. European Commission (1999a) Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions on certain Community measures to combat discrimination, COM (1999) 564 final, Brussels, European Commission.

70. European Commission (1999b) Implementing the framework for financial markets: action plan, COM(1999) 232 final, Brussels.

71. European Commission (2000) Data protection: Commission adopts decisions recognising adequacy of regimes in US, Switzerland and Hungary, IP/00/865, Brussels.

72. European Commission (2002) Proposal for a Directive of the European Parliament and of the Council on the harmonisation of the laws, regulations and administrative provisions of the Member States concerning credit for consumers, COM(2002) 443 final, Brussels.

73. European Commission (2003) Internal Market Scoreboard, Nr. 12, 5 May 2003, Brussels, http://europa.eu.int/comm/internal_market/en/update/score/index.htm.

74. European Parliament (1995) Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, *Official Journal of the European Communities*, L 281, 0031 - 0050.

75. European Parliament (1997) Directive 97/5/EC of the European Parliament and of the Council of 27 January 1997 on cross-border credit transfers, *Official Journal of the European Communities*, L 043, 0025 - 0030.

76. European Parliament (1998) Directive 98/7/EC of the European Parliament and of the Council of 16 February 1998 amending Directive 87/102/EEC for the approximation of the laws, regulations and administrative provisions of the Member States concerning consumer credit, *Official Journal of the European Communities*, L 101, 0017 - 0023.

77. European Parliament (2000) Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce'), *Official Journal of the European Communities*, L 178, 0001 - 0016, http://europa.eu.int/eur-lex/pri/en/oj/dat/2000/l_178/l_17820000717en00010016.pdf.

78. Eurostat (2003) Eurostat Yearbook 2003, http://europa.eu.int/comm/eurostat/.

79. Farewell, V.T. and Prentice, R.L. (1980) The approximation of partial likelihood with emphasis on case-control studies, *Biometrika,* 67, 273 - 278.

80. FCRA (1970) The Fair Credit Reporting Act, http://www.ftc.gov/privacy/fcract/.

81. Fox, J. (1991) *Regression Diagnostics*, Sage.

82. Freund, R.J. and Littel, R.C. (2000) *SAS system for regression*, SAS Institute Inc., Cary, NC, USA.

83. GLBA (1999) The Financial Modernization (Gramm-Leach Bliley) Act, http://www.ftc.gov/privacy/glbact/.

84. Goode, R.M. (1989) *Consumer credit law*, London : Butterworths.

85. Goode, R.M. and Cranston, R. (1993) *Commercial and consumer law : national and international dimensions*, Oxford : Clarendon Press.

86. Guardia, N.D. (2000) Consumer Credit in the European Union, ECRI Research Report No 1, http://www.ecri.be/.

87. Hand, D.J. (1979) Psychiatric examples of Simpson's paradox, *British Journal of Psychiatry*, 135, 90 - 91.

88. Hand, D.J. and Henley, W.E. (1993) Can reject inference ever work?, *IMA Journal of Mathematics Applied in Business & Industry*, 5, 45 - 55.

89. Hand, D.J. and Henley, W.E. (1996) Some developments in statistical credit scoring, in *Machine learning and statistics: the interface*, ed. Nakhaeizadeh, G. and Taylor, C. C., John Wiley & Sons, Chichester.

90. Hand, D.J. (1997) *Construction and Assessment of Classification Rules*, John Wiley & Sons, Chichester.

91. Hand, D.J. (1998) Consumer credit and statistics, in *Statistics in Finance*, ed. Hand, D. J. and Jacka, S. D., Edward Arnold, London.

92. Hand, D.J. and Adams, N.M. (2000) Defining attributes for scorecard construction in credit scoring, *Journal of Applied Statistic*, 27, 5, 527 - 540.

93. Hand, D.J. and Kelly, M.G. (2001) Lookahead scorecards for new fixed term credit, *Journal of Operational Research Society*, 52, 989 - 996.

94. Hand, D.J. (2003) Good practice in retail credit score card assessment, in *Credit Scoring and Credit Control VIII Conference Proceedings*, Credit Research Centre, The University of Edinburgh.

95. Hand, D.J. and Vinciotti, V. (2003) Building scorecards for today's customers, in *Credit Scoring and Credit Control VIII*, Credit Research Centre, University of Edinburgh.

96. Hanley, J.A. and McNeil, B.J. (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology,* 143, 1, 29 - 36.

97. Hanley, J.A. and Theriault, G. (2000) Simpson's paradox in meta-analysis, *Epidemiology*, 11, 5, 613.

98. Harrison, G.W. (1998) Mortgage Lending in Boston: A Reconsideration of the Evidence, *Economic Inquiry*, 36, 1, 29 - 38.

99. Henley, W.E. (1995) Statistical aspects of credit scoring, PhD Thesis, The Open University, Milton Keynes, UK.

100. Henley, W.E. and Hand, D.J. (1996) A k-NN classifier for assessing consumer credit risk, *The Statistician*, 45, 77 - 95.

101. Hirschman, E.C. (1982) Consumer payment systems: the relationship of attribute structure to preference and usage, *Journal of Business*, 55, 531 - 545.

102. Hopper, M.A. and Lewis, E.M. (1992) Development and use of credit profit measures for account management, *IMA Journal of Mathematics Applied in Business and Industry*, 4, 3 - 17.

103. Hosmer, D.W. and Lemeshow, S. (2000) *Applied Logistic Regression*, 2nd edition, John Wiley & Sons.

104. Howells, G. (1995) Data protection, confidentiality, unfair contract terms, consumer protection and credit reference agencies, *Journal of Business Law*, 44, 2.

105. HMDA (1975) Regulation C - Home Mortgage Disclosure, *12 U S Code*, 2801 et seq., http://www.ffiec.gov/hmda/pdf/regulationc2004.pdf.

106. HRA (1998) Human Rights Act, http://www.hmso.gov.uk/acts/acts1998/19980042.htm.

107. Hsia, D.C. (1978) Credit Scoring and the Equal Credit Opportunity Act, *The Hastings Law Journal*, 30, 371 - 448.

108. Japelli, T. and Pagano, M. (2000) Information Sharing, Lending and Defaults: Cross-Country Evidence, Working Paper No.22, University of Salerno.

109. Jentzsch, N. (2003). The Implications of the New Consumer Credit Directive for EU Market Integration, Testimony before the Committee on Legal Affaires and the Internal Market, European Parliament on April 29, 2003, http://www.europarl.eu.int/hearings/20030429/juri/jentzsch1_en.pdf

110. Joanes, D.N. (1993) Reject inference applied to logistic regression for credit scoring, *IMA Journal of Mathematics Applied in Business & Industry*, 5, 35 - 43.

111. Johnson, R.W. (1992) Legal, Social and Economic Issues in Implementing Scoring in the US, in *Credit Scoring and Credit Control*, ed. Thomas, L. C., Crook, J. N., and Edelman D.B., Clarendon Press, Oxford.

112. Julious, S.A. and Mullee, M.A. (1994) Confounding and Simpson's paradox, *British Medical Journal*, 309, 6967, 1480 - 1481.

113. Kalbfleisch, J.D. and Prentice, R.L. (1980) *The Statistical Analysis of Failure Time Data*, John Wiley & Sons.

114. Kaplan, E.L. and Meier, P. (1958) Non-parametric estimation from incomplete observations, *Journal of American Statistical Association*, 53, 457 - 481.

115. Kelly M.G., Hand, D.J., and Adams, N.M. (1999) The impact of changing populations on classifier performance, in *Knowledge Discovery and Data Mining KDD-99*, ed. Chaudhuri, S. and Madigan, D., ACM Press, San Diego CA USA.

116. Koopman, J. (1999) Policy and practice in the Netherlands, *Journal of Consumer Studies and Home Economics*, 23, 2, 89-101.

117. Korff, D. (1998) *The feasibility of a seamless system of data protection rules for the European Union*, European Commission, Brussels.

118. Krastanova, A. (2003) ECRI Statistical Package: Consumer Credit and Lending in the EU and the CEECs for the period 1990-2002, ECRI, http://www.ecri.be/statistics.html.

119. Lawless, J.F. (1981) *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons.

120. Le Conseil d'Etat (2001) Statuant au contentieux, sur le rapport de 10ème sous-section – N° 204909 - Séance du 19 octobre 2001, lecture du 30 octobre 2001, http://www.conseil-etat.fr/ce-data/index2.htm.

121. Leonard, K.J. (1997) Behavior Scores to Predict Profitability, Credit Scoring and Credit Control V Conference Proceedings, Credit Research Centre, The University of Edinburgh.

122. Lewis, E.M. (1992a) *An Introduction to Credit Scoring*, Athena Press, San Rafael.

123. Lewis, E.M. (1992b) Credit scoring and credit control from four points of view, in *Credit Scoring and Credit Control*, ed. Thomas, L. C., Crook, J. N., and Edelman D.B., Clarendon Press, Oxford.

124. Liebowitz, S.J. (1993) A Study that Deserves No Credit, *Wall Street Journal*, September 1, 1993, p.A14

125. Lindley, J.T., Rudolf, P., and Selby, E.B. (1989) Credit card possession and use: change over time, *Journal of Economics and Business*, 41 , 127 - 142.

126. Lovie, A.D. and Lovie, P. (1986) The flat maximum effect and linear scoring models for prediction, *Journal of Forecasting*, 5, 3, 159 - 168.

127. Lucas, A. (1992) Updating scorecard: removing the mystique, in *Credit Scoring and Credit Control*, ed. Thomas, L. C., Crook, J. N., and Edelman, D. B., Oxford University Press, Oxford.

128. Mahlman, M. (2003) Executive summary on race equality directive. State of play in Germany, http://europa.eu.int/comm/employment_social/fundamental_rights/legis/mslegl nracequal_en.htm.

129. Makuch, W.M. (1998) The basics of a better application score, in *Credit Risk Modeling*, ed. Mays, E., Glenlake Publishing Company, LTD; Fitzroy Dearborn Publishers, Chicago.

130. Mannila, H. (1995) Aspects of data mining, in *Statistics, Machine Learning and Knowledge Discovery in Databases*, ed. Kodratoff, Y., Nakhaeizadeh, G., and Taylor, C. C., Heraklion, Crete.

131. Martin, R.E. and Hill, R.C. (2000) Loan performance and race, *Economic Inquiry*, 38, 136 - 150.

132. Mester, L.J. (1997) What's the point of credit scoring?, *Business Review*, 5.

133. Munnel, A.H., Tootell, G.M.B., Browne, L.E., and McEneaney, J. (1992) Mortgage lending in Boston: Interpreting HMDA Data, Working Paper, 92-7, Federal Reserve Bank of Boston.

134. Munnel, A.H., Tootell, G.M.B., Browne, L.E., and McEneaney, J. (1996) Mortgage Lending in Boston: Interpreting HMDA Data, *The American Economic Review*, 86 , 1, 25 - 53.

135. Nakhaeizadeh, G., Taylor, C.C., and Kunisch, G. (1996) Dynamic aspects of statistical classification, in *Intelligent Adaptive Agents, AAAI Technical report No. WS-96-04*, AAAI Press, Menlo Park, CA.

136. Nakhaeizadeh, G., Taylor, C.C., and Kunisch, G. (1997) Dynamic supervised learning: some basic issues and applicaton aspects, in *Classification, Data Analysis and Knowledge Organisation*, ed. Klar, B. and Opitz, O., Springer Verlag.

137. Narain, B. (1992) Survival analysis and the credit granting decision, in *Credit Scoring and Credit Control*, ed. Thomas, L. C., Crook, J. N., and Edelman, D. B., Oxford University Press, Oxford.

138. Nevin, J.R. and Churchill, G.A., Jr. (1979) The Equal Credit Opportunity Act: An Evaluation, *Journal of Marketing*, 43, 2, 95 - 104.

139. Office of Comptroller of Currency (1994) Bulletin 94-30. http://www.occ.treas.gov/ftp/bulletin/97-24att.pdf

140. Office of Comptroller of Currency (1997) Bulletin 97-24. http://www.occ.treas.gov/ftp/bulletin/97-24att.pdf

141. OFT (1992) Credit scoring: a report by Director General by Fair Trading, Office of Fair Trading.

142. OFT (2000) *Guide to Credit Scoring*, http://www.bba.org.uk/public/consumers.

143. Overstreet, G.A.J., Bradley, E.L., and Kemp, R.S. (1992) The flat-maximum effect and generic linear scoring models: a test, *IMA Journal of Mathematics Applied in Business & Industry*, 4, 1, 97 - 109.

144. Overstreet, G.A.J. and Bradley, E.L.J. (1996) Applicability of generic linear scoring models in the U.S. credit-union environment, *IMA Journal of Mathematics Applied in Business & Industry*, 7, 291 - 311.

145. Oxford English Dictionary (1933) *Volume III.*, At the Clarendon Press, Oxford.

146. Oxley, J. (2003) Credit limit strategy optimisation, in *Credit Scoring and Credit Control VIII*, Credit Research Centre University of Edinburgh.

147. Parmar, M.K.B. and Machin, D. (1995) *Survival analysis: a practical approach*, John Wiley & Sons, Chichester.

148. Phelps, E.S. (1972) The Statistical Theory of Racism and Sexism, *American Economic Review*, 62, 9, 659 - 661.

149. Platts, G. and Howe, I. (1997) A single European scorecard, in *Credit Scoring and Credit Control V, Conference Proceedings*, Management School, Edinburgh.

150. Peterson, R.L. and Peterson, C.M. (1978) Testing for Sex Discrimination in Commercial Bank Consumer Lending, Working Paper, 10, West Lafayette, IN: Purdue University, Credit Research Center.

151. Prentice, R.L., Williams, B.J., and Peterson, A.V. (1981) On the regression analysis of multivariate failure time data, *Biometrika*, 68, 373 - 379.

152. Rapoport, A. (1975) Research paradigms for studying dynamic decision behaviour, in *Utility, Probability and Human Decision Making*, ed. Wendt, D. and Vlek, C. A. J., Reidel, Dordrecht.

153. Regulations (2003) *Statutory Instrument No. 1626. The Race Relations Act 1976 (Amendment) Regulations 2003*, http://www.hmso.gov.uk/si/si2003/20031626.htm.

154. RRA (1976) *Race Relations Act*, http://www.homeoffice.gov.uk/docs/part1.html.

155. Schebesch, K.B. and Stecking, R. (2003) Support vector machines for classifying and describing credit applicants: detecting typical and critical regions, in *Credit Scoring and Credit Control VIII*, Credit Research Centre The University of Edinburgh.

156. Sen, P.K. (1960) On some convergence properties of U-statistics, *Calcutta Statistical Association Bulletin*, 10, 1, 18.

157. SDA (1975) The Sex Discrimination Act, http://www.pfc.org.uk/legal/sda.htm.

158. Shay, R.S. and Genderton, D.E. (1979) Anti-Discrimination Laws in Consumer Credit Markets: Their Impact on Creditors Approval of Applications, in *The Costs and Benefits of Public Regulation of Consumer Financial Services*, ed. Heggestad A.A. and Mingo J.J., Cambridge, MA: Abt Associates, Inc.

159. Shinkel, B.A. (1980) The Effects of Equal Credit Opportunity Legislation in Consumer Finance Lending, *Journal of Business Research*, 8, 3, 113 - 134.

160. Simpson, E.H. (1951) The interpretation of interaction in contingency tables, *Journal of the Royal Statistical Society B*, 13, 238 - 241.

161. Singleton, S. (1999) Privacy and Human Rights: Comparing the United States to Europe', CATO Institute, http://www.cato.org/pubs/wtpapers/991201paper.html

162. Srinivasan, V. and Kim, Y.H. (1987) Credit granting: a comparative analysis of classification procedures, *The Journal of Finance*, 42, 665 - 683.

163. Stablein, D.M., Carter, W.H., Jr., and Novak, J.W. (1981) Analysis of survival data with nonproportional hazard functions, *Controlled Clinical Trials*, 2, 149 - 159.

164. Staten, M.E. (1999) Strength in numbers: aggregate credit bureau data benefit portfolio management, *The Journal of Lending & Credit Risk Management*, 82, 1, 27 - 31.

165. Stepanova, M. (2001) Using survival analysis methods to build credit scoring models, PhD Thesis, University of Southampton.

166. Stepanova, M. and Thomas, L.C. (2001) PHAB scores: proportional hazards analysis behavioural scores, *Journal of Operational Research Society*, 52, 1007 - 1016.

167. Stepanova, M. and Thomas, L.C. (2002) Survival analysis methods for personal loan data, *Operations Research*, 50, 2, 277 - 289.

168. Taylor, C.C. and Nakhaeizadeh, G. (1997) Learning in dynamically changing domains: theory revision and context dependence issues, in *Machine learning: ECML-97; 9th European Conference on Machine Learning* , Springer, Prague.

169. Theil, H. (1971) *Principles of Econometrics*, John Wiley, New York.

170. Therneau, T.M. and Grambsch, P.M. (2000) *Modelling Survival Data*, Springer-Verlag.

171. Thomas, L.C. (1998) Methodologies for classifying applicants for credit, in *Statistics in Finance*, ed. Hand, D. J. and Jacka, S. D., Edward Arnold, London.

172. Thomas, L.C. (2000) A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers, *International Journal of Forecasting*, 16, 2, 149 - 172.

173. Thomas, L.C., Edelman, D.B., and Crook, J.N. (2002) *Credit scoring and its applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

174. Thomas, L.C., Thomas, S., Tang, L., and Gwilym, O.A. (2003) The impact of demographic and economic variables on finacial policy purchase timing decisions, in *Credit Scoring and Credit Control VIII*, Credit Research Centre University of Edinburgh.

175. Till, R.J. (2001) Predictive behavioural models in credit scoring and retail banking, PhD Thesis, Department of Mathematics, Imperial College of Science, Technology and Medicine.

176. Tootell, G.M.B. (1996) Redlining in Boston: Do Mortgage Lenders Discriminate Against Neighborhoods?, Federal Reserve Bank of Boston, Working Paper No.96-6.

177. Treaty (1997) Treaty of Amsterdam, *Official Journal of the European Communities*, C 340.

178. Treaty (2002) Consolidated version of the Treaty establishing the European Community, *Official Journal of the European Communities*, C325 , 0033 - 0184, http://europa.eu.int/eur-lex/en/treaties/dat/EC_consol.pdf.

179. Urwin, D.W. (1995) *The Comminity of Europe: A History of European Integration since 1945*, Longman.

180. Usher J. A. (1998) *General Principles of EC Law*, Longman.

181. Usher J. A. (2000) *Law of money and financial services in the EC* , Longman.

182. Van Order, R., Westin, A. M., and Zorn, P. (1993), "*Effects of Racial Composition of Neighborhoods on Defaults, and Implication in Mortgage Markets*", Paper presented at the Allied Social Science Associations meetings in Anaheim, CA.

183. Volker, P.A. (1982) A note on factors influencing the itilization of bankcard, *Economic Records*, September , 281 - 289.

184. von Winterfeldt, D. and Edwards, W. (1982) Costs and payoffs in perceptual research, *Psychological Bulletin*, 91, 609 - 622.

185. Wachter, S. (1997) Discrimination in Financial Services: What Do We Know?, *Journal of Financial Services Research* , 11, 1, 205.

186. Weatherill, S. (1997) *EC Consumer law and Policy*, Longman.

187. Wei, L.J., Ying, Z., and Lin, D.Y. (2003) Linear regression analysis of censored survival data based on rank tests, *Biometrika*, 77, 845 - 851.

188. Westin, A.M. (1967) *Privacy and freedom*, Atheneum, New York.

189. White, K.J. (1975) Consumer choice and the use of bank credit cards: a model and cross-section results, *Journal of Consumer Research*, 2, 10 - 18.

190. Widmer, G. and Kubat, M. (1993) Effective learning in dynamic environments by explicit context tracking, in *Proceedings of the Sixth European Conference on Machine Learning*, Springer-Verlag, Berlin.

191. Widmer, G. (1994) Combining robustness and flexibility in learning drifting concepts, in *Proceedings of the Eleventh European Conference on Artificial Intelligence*, Wiley, Chichester.

192. Widmer, G. (1997) Tracking context changes through meta-learning, *Machine Learning*, 27 , 3, 259 - 287.

193. Wilkinson, G. (1992) How credit scoring really works, in *Credit Scoring and Credit Control*, ed. Thomas, L. C., Crook, J. N., and Edelman, D. B., Oxford University Press, Oxford.

194. Yinger, J. (1997) Discrimination in Mortgage Lending: A Literature Review, in *Mortgage lending, racial discrimination, and federal policy*, ed. Goering J.M. and Wienk, R. E., Aldershot: Ashgate.

195. Yobas, M.B., Crook, J.N., and Ross, P. (1997) Credit scoring using neural and evolutionary techniques, Working Paper Series, 97/2, Credit Research Centre, University of Edinburgh.

196. Yule, G.U. (1903) Notes on theory of association of attributes in statistics, *Biometrica*, 2, 121 - 134.

197. Zandi, M. (1993) Boston Fed's Bias Study Was Deeply Flawed, *The American Banker*, August 19, 1993.

198. Zwamborn, M. (2003) Executive summary on race equality directive. State of play in the Netherlands., http://europa.eu.int/comm/employment_social/fundamental_rights/legis/mslegl nracequal_en.htm.

APPENDIX A

Table A1. Characteristics available for analysis

	Characteristic	Belgium	Netherlands	Germany
	Application			
1	**Telephone**	X	X	X
2	**Residential status**	X	X	X
3	**Marital status**	X	X	X
4	**Occupation**	X	X	X
5	**Age**	X	X	X
6	**Time at address since 18**	X	X	X
7	**Time in employment**	X	X	X
8	Time at bank since 18	X		X
9	**Type of business**	X	X	X
10	Occupation of spouse	X	X	
11	Credit card type	X		X
12	Agreement type	X	X	X
13	**No of children/dependants**	X	X	X
15	**Spouse age**	X	X	X
16	**Goods code**	X	X	X
17	Identity	X		X
18	Language	X		
19	Postcode current address	X	X	X
20	Total addresses given	X		
21	**Employer's phone**	X	X	X
22	**Card insurance**	X	X	X
23	**Credit insurance**	X	X	X
24	Product insurance			X
25	**Goods price**	X	X	X
26	Initial instalment	X		X
28	**Payment date**	X	X	X
29	Fast decision	X	X	
31	Section number		X	
32	Dealer number		X	
33	Original balance		X	
34	Override code	X		X
35	Underwriter	X		X
36	Retailer	X	X	
37	Time in benefit		X	
38	Type of bank account		X	
39	Bank name		X	
40	Nationality		X	X
41	Bank sort code	X		X
42	Postcode previous address	X		

43	Time at previous address	X		X
	Characteristic	**Belgium**	**Netherlands**	**Germany**
44	Spouse indicator (yes, no)	X		X
45	2-nd cardholder			X
46	Time at previous job	X		X
47	Deposit		X	X
48	% deposit		X	X
49	Requested credit limit		X	X
50	Given credit limit		X	X
51	Credit limit live revolving		X	
52	Credit limit live store cards		X	
53	Loan amount		X	X
54	Credit as % of goods price		X	
55	Term of loan		X	
56	Income 1st applicant		X	
57	Income co-app		X	
58	Mortgage/rent		X	
59	Other costs		X	
60	Instalment paid		X	
61	Type of service (phone, budget, etc.)		X	
62	Type of bank (normal, post, giro)			X
63	Card status (live, lost, stolen, etc.)		X	X
64	East/West indicator			X
	Bureau			
65	Total contracts	X		X
66	Time first registration	X	X	X
67	Time since last registration/most recent account (H)/	X		X
68	Total contracts normal balance	X		
69	Total contracts in arrears	X		
70	Total amount of credit/credits ever	X		X
71	Total amount of liability	X		
72	Total amount of reduction	X		
73	Mortgage loan marker	X		
74	No of live cards + revolving		X	
75	No of paid-up cards + revolving		X	
76	No of paid fixed term accounts		X	
77	Bureau negative		X	X
78	Number of live fixed term accounts		X	
79	Total A accounts		X	
80	Total A+ accounts		X	
81	Total outstanding balances		X	
82	Known at bureau		X	
83	Total A accounts opened		X	
84	Total A accounts closed		X	
85	Total A+ accounts opened		X	

86	Total A+ accounts closed		X	
	Characteristic	**Belgium**	**Netherlands**	**Germany**
87	Time since last bad A opened		X	
88	Time since last bad A+ opened		X	
91	No of live HP		X	
92	No of live revolving accounts		X	
93	No of live card accounts		X	
94	No of live extensions		X	
95	No of live overdrafts		X	
96	No of live loan accounts		X	
97	No of live mortgage accounts		X	
98	No of live other accounts		X	
99	No of good paid up HP		X	
100	No of good paid up revolving accounts		X	
101	No of good paid up card accounts		X	
102	No of good paid up extensions		X	
103	No of good paid up overdrafts		X	
104	No of good paid up loan accounts		X	
105	No of good paid up mortgage accounts		X	
106	No of good paid up other accounts		X	
107	Outstanding balance HP accounts		X	
108	Outstanding balance live extensions		X	
109	Outstanding balance live overdraft accounts		X	
110	Outstanding balance live loan accounts		X	
111	Total live bureau		X	
112	Total closed bureau		X	
113	Last bad A		X	
114	Last bad A+		X	
115	No new accounts in last 6 months		X	
116	Bureau time known			X
117	Credit enquiries last 10 days			X
118	Bureau quantity 1			X
119	Bureau quantity 2			X
120	Bureau quantity 3			X
121	Bureau quantity 4			X
123	Bureau amount			X
124	Bureau low process			X
125	Bureau major negative			X
126	Bureau minor negative			X
128	Total no credits still open			X
129	Total amount credits repaid			X
130	Bureau score			X

Table A2. Attribute coding by country

Characteristics	Attributes	Belgium	Netherlands	Germany	Generic
Telephone	NOT GIVEN	0,N	3	0	0
	GIVEN	Y,1	1	1	1
	SECRET		2	2	2
	MOBILE	3	4	3	3
	NO INFO	" ", X	" "	" "	" "
Residential status	HOMEOWNER	1	02	1	1
	RENTED HOUSE	2	01	2	2
	RENTED FLAT	3	03	3	3
	RENTED ROOM	4	07	4	4
	LIV w PARENT	5	04	5	5
	CARAVAN	6	05	6	6
	HOUSE BOAT		06		7
	OTHER		99		99
	NO INFO	" "	" "	" "	" "
Marital status	MARRIED	1	1	1	1
	SINGLE	2	2	2	2
	DIVORCED	3	3	3	3
	WIDOW	4	5	4	4
	LIVING TOG	5	4	5	5
	LIV TOG REG		6		6
	NO INFO		" "		" "
Time at address	with Parents	9999			9999
Time in employment	No Info			.	9991
	Part-time		9992		9992
	Military		9993		9993
	Benefit Work		9994		9994
	Agency		9995		9995
	Retired	9996	9996	9996	9996
	Housewife	9997	9997	9997	9997
	Student	9998	9998	9998	9998
	Jobless			9995	9999
	Allowance	9999			9999
Spouse age	No spouse	99	999	999	999

Characteristics	Attributes	Belgium	Netherlands	Germany	Generic
Occupation	Employed/Full-time	01	01	01	1
	Self Employed	03	02	02	2
	Retired	05		03	3
	Housewife	04	20	04	4
	Student	06	05	05	5
	Military Service		06	06	6
	Allowance 1	07		07	7
	Allowance 2	08		08	8
	Allowance 3	09		09	9
	Arb prv mit			10	10
	Witwen rente			11	11
	Benefit		21		20
	Ben Wrk <50		09		21
	Ben Wrk >50		10		22
	PT	02			30
	PT 0-8 Hrs		11		31
	PT 9-32 Hrs		12		32
	PT 33-40 Hrs		13		33
	Agency	00			40
	Agency < 1Y		07		41
	Agency > 1Y		08		42
	No Info	0	" "	" "	" "
Credit insurance	No Info	" "	" "	" "	" "
	No insurance	0	0	0, N	0
	Credit insurance	1	1	1	1
	Other	2	3		2
			4		2
			5		2
			6		2
			7		2
			8		8
			9		2
Card insurance	No Info	" "	" "	" "	" "
	No insurance	0	0	0	0
	Card insurance	1	1	1	1
	Other	2	2	2	2

Characteristics	Attributes	Belgium	Netherlands	Germany	Generic
Type of business	Industry	01	01	09	1
	Officials	02	02		2
	Building Ind	03	03	12	3
	Pub Trs/Post	04	04		4
	Prof Soldier	05	05		5
	Health Care	06	06	11	6
	Education	07	07		7
	Bank/insurance	08	08	02	8
	Catering	09	09	06	9
	Shop Employee	10	10	10	10
	Cleaning Agc	11	11		11
	Agrarian Sec	12	12	08	12
	Harbour Ind	13	13		13
	Road Transport	14	14		14
	Shipping Ind	15	15		15
	Aviation	16	16		16
	Craftsman	17	17		17
	Business	18	18		18
	Service prof	19	19		19
	Computer Ind	20	20		20
	Aus/Erz/Lehr			04	27
	High Tech			03	20
	Emp Gov Sub		21		21
	Offentlichtr			01	22
	Sonstge Univ			05	23
	Media			07	24
	Unv-Abschlss			13	25
	Dienstleist.			14	26
	21	21			31
	Benefit AAW		51		51
	Benefit ABW		52		52
	Benefit AOW		53		53
	Benefit AWW		54		54
	Benefit RWW		55		55
	Benefit WAO		56		56
	Benefit WW		57		57
	Benefit WWW		58		58

Characteristics	Attributes	Belgium	Netherlands	Germany	Generic
Type of business	Benefit ROA		59		59
	Benefit VUT		60		60
	Militry Serv		61		61
	Student+Job		62		62
	Agency		63		63
	Pensioner		64		64
	Indep Means		65		65
	Housewife		66		66
	Unknown	99	99	25	99
Employer's phone	No Info	" "		" "	" "
	No Info	" "	" "	" "	" "
	Not Given	0	2	0	0
	Given	1	1	1	1
Goods price	Secret number			2	2
Goods code		Belgian Francs	Dutch Guilders	German Marks	EUR
	CARD APPLICA	00	00	CA	0
	MONO TV	01			1
	COLOUR TV	02			1
	TV		01	01	1
	VIDEO RECORD	03	02	02	3
	TV/VIDEO COM	04	08	03	4
	HIFI RADIO	05	03	05	5
	MICROWAVE	06	16	11	6
	COOKER	07	15	14	7
	FRIDGE	08	10	12	8
	WASHING MACHINE	09	11	13	9
	VACUUM CLNR	10	17		10
	DISHWASHER	11	14		11
	HEATER	12			12
	DRIER	13	12		13
	VIDEO CAMERA	14	21		14
	CD PLAYER	15	04		15
	FREEZER	16	13		16
	HSEHLD GOODS	17	18	19	17

Characteristics	Attributes	Belgium	Netherlands	Germany	Generic
	VID/CAS TAPE	18	07		18
	COMPACT DISC	32	07		18
	FAX		32		19
	PHONE FAX	19		25	19
	MOBILFUNK			24	19
	TELEPHONE		30		19
	OTHR PHONES	20	33		20
	CAR PHONE		31		20
	21	21			21
	22	22			22
	23	23			23
	24	24			24
	OTHERS	25	19		25
	BEDROOM FURN	26	40	33	26
	DINING FURN	27	41		26
	LIVING FURN	28	42	34	26
	KITCHEN FURN	29		36	26
	FIRNITURE		49	30	26
	BEDDING	30		61	30
	SUNBEDS	31	23		31
	33	33			33
	34	34			34
	35	35			35
	36	36			36
	OTHERS	37			37
	CARPETS	38	44	32	38
	MENS CLOTHES	39			39
	BOYS WEAR	40			40
	MENS BOUTIQUE	41			41
	OTHERS	42			42
	GARDEN FURN	43	46	35	43
	LAWN MOWERS	44			44
	OTHERS	45			45

Characteristics	Attributes	Belgium	Netherlands	Germany	Generic
Goods code	PIANOS	46	60		46
	ORGANS	47	61		47
	OTHER INSTRUMENTS	48	62		48
	CAMERAS	49		21	49
	NEVER USED	50			50
	SCANNER	51		22	51
	MENS SHOES	52			52
	LADIES SHOES	53			53
	CHILD SHOES	54			54
	SPORTS GOODS	55			55
	NEW MOTORCYC	56	52		56
	MEN/CHL FASH	57			57
	MOPEDS	58	50		58
	CYCLES	59	51	41	59
	LADIES CLOTHES	60	71		60
	GIRLS CLOTHES	61			61
	CHILD CLTHS		72		61
	OTHERS	62	79		62
	MENS CLOTHES	63	70		63
	LADY BOUTIQU	65	82		65
	COMPUTERS	66	27	20	66
	DBLE GLAZING	67			67
	BLTIN KITCHN	68	65		68
	BBY/CHLD FURN	69			69
	SOCKS	70			70
	BEACHWEAR	71	77		71
	LINGERIE	72	76		72
	IND/CLOTHES	73			73
	HSEHLD TEXTL	74			74

Characteristics	Attributes	Belgium	Netherlands	Germany	Generic
Goods code	KNIT MACHINE	75			75
	DRESS MATERL	76			76
	COMPUTER REQ	77	28		77
	JEWELLERY	78	83		78
	79	79			79
	STONE CLAD.	80			80
	81	81			81
	BLDG MATLS	82	67	50	82
	83	83			83
	PARFUMERIE	84	81		84
	CAR RADIOS	85	06		85
	SPORTS ARTIC	86			86
	LEATHERWEA R	87	80		87
	RESTAURANT	88			88
	MISC OTHERS	89	89		89
	CARAVANS	90			90
	FRONT TENTS	91			91
	92	92			92
	93	93			93
	94	94	91		94
	95	95	94		95
	96	96	95		96
	CODE 97	97	97		97
	NEVER USED	98			98
	OTHERS	99		90	99
	LAMPS		05	31	101
	STOVE		36		102
	MISC BROWN G		09		103
	SATELLITE ANT			23	104
	MISC ELECT		29		110
	GARDEN/MISC			53	111
	GARDEN MATER		66		112

Characteristics	**Attributes**	**Belgium**	**Netherlands**	**Germany**	**Generic**
Goods code	CURTAINS		45	60	113
	PHOTO EQUIP		20	26	114
	MISC PHOTOS		22		115
	PRINTER			27	116
	MONITOR			28	117
	GAME COMPUTR		25,26		118
	OFFICE MAC		34		120
	OFFICE EQUIP		35		121
	MALER BEDARF			51	123
	TOOLS		68	52	124
	MISC DIY		69		124
	CAR EQUIP		64	40	126
	CAMPING EQUP		85		127
	BLUMEN			91	128
	NO INFO			" "	" "

Table A3. Coarse-classification by country. Home telephone.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
Given	0.03	1	Given	0.16	1	Not given	-0.80	1
Not given	-0.55	2	Secret	0.24	1	Given	0.07	2
Mobile	-0.48	3	Not given	-0.67	2	Secret	-0.32	2
			Mobile	-1.00	2			

Table A4. Coarse-classification by country. Employer's telephone.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
Not given	-4.05	1	Not given	0.00	1	Not given	0.64	1
Given	-4.01	2	Given	-1.00	2	Given	-0.06	2

Table A5. Coarse-classification by country. Residential status.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
HOME-OWNER	0.55	1	HOME-OWNER	0.81	1	HOME-OWNER	0.88	1
CARAVAN	0.18	2	RENTED HOUSE	-0.08	2	RENTED HOUSE	0.18	2
RENTED HOUSE	-0.18	2	RENTED FLAT	-0.20	3	CARA-VAN	0.36	2
RENTED FLAT	-0.32	3	HOUSE BOAT	0.15	3	RENTED FLAT	-0.05	3
RENTED ROOM	-0.95	4	OTHER	-0.14	3	RENTED ROOM	-0.86	4
LIV w PARENT	-0.56	4	LIV w PARENT	-0.63	4	LIV w PARENT	-0.64	4
			CARA-VAN	-0.97	5			
			RENTED ROOM	-0.85	5			

Table A6. Coarse-classification by country. Credit insurance.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
No insurance	-4.06	1	No insurance	-0.03	1	No insurance	-0.43	1
Credit insurance	-3.75	2	Credit insurance	-0.09	2	Credit insurance	0.01	2
Other	-3.51	2	Other 4	-0.59	2			
			Other 6	-0.02	2			
			Other 8	0.13	3			
			Other 9	-0.27	2			

Table A7. Coarse-classification by country. Card insurance.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
No insurance	-4.03	1	No insurance	0.05	1	No insurance	0.04	1
Card insurance	-4.03	2	Card insurance	-0.22	2	Card insurance	-0.34	2
			Other	-0.33	2			

Table A8. Coarse-classification by country. Occupation.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
Empl-d/ Full-time	0.97	1	Empl-d/ Full-time	-0.04	3	Retired	0.86	1
PT	1.10	2	Self Empl-d	-0.31	1	Allow-ance 2	-0.41	2
House-wife	1.11	2	Agency < 1Y	-0.63	2	Allow-ance 1	0.12	2
Self Empl-d	0.72	3	Agency > 1Y	-0.91	2	Empl-d/ Full-time	-0.05	3
Allow-ance 3	1.72	4	Ben Wrk <50	0.37	4	Student	-0.03	3
Retired	2.29	4	PT 0-8 Hrs	0.91	4	House-wife	-0.02	3
Allow-ance 1	0.87	5	PT 9-32 Hrs	0.39	4	Allow-ance 3	-1.32	5
Allow-ance 2	0.93	5	PT 33-40 Hrs	0.20	4	Military Ser	-0.50	5
Agency	0.45	6	House-wife	0.47	4	Self Empl-d	-0.33	4
			Benefit	0.23	4			

Table A9. Coarse-classification by country. Type of business.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
Unknown	0.19	0	Agency	-0.76	1	Catering	-0.64	1
Harbour Ind	-0.56	1	Prof Soldier	-0.44	2	Sonstge Univ	-0.32	2
Catering	-0.54	1	Catering	-0.36	2	Media	-0.29	2
Craftsman	-0.45	1	Benefit RWW	-0.30	2	Agrarian Sec	-0.32	2
Cleaning Agc	-0.45	1	Benefit AAW	-0.29	2	Building Ind	-0.39	2
Agrarian Sec	-0.39	1	Agrarian Sec	-0.28	2	Aus/Erz/L ehr	-0.20	3
Building Ind	-0.38	1	Student+ Job	-0.24	2	Shop Emplyee	-0.10	3
Road Transpt	-0.32	2	Building Ind	-0.18	3	Dienstleist .	-0.16	3
Shop Emplyee	-0.21	2	Craftsma n	-0.14	3	Industry	0.05	4
Business	-0.14	3	Shop Emplyee	-0.13	3	Unv-Abschlss	0.01	4
Industry	-0.13	3	Industry	-0.12	3	Offentlicht r	0.28	5
Aviation	-0.13	3	Road Transpt	-0.11	3	Bank/insur ce	0.37	5
Computer Ind	-0.06	4	Cleaning Agc	-0.08	4	High Tech	0.48	6
Service prof	0.00	4	Business	-0.03	4	Health Care	0.44	6
21	0.08	4	Unknow n	0.00	4	No info	0.62	7
Pub Trs/Post	0.18	5	Service prof	0.09	4	15	0.50	7
Shipping Ind	0.18	5	Compu-ter Ind	0.19	4	16	0.42	7
Officials	0.24	5	Bank/ins urace	0.22	4	Unknown	0.76	7
Health Care	0.27	5	Aviation	0.24	4			
Bank/insu race	0.46	6	Pub Trs/Post	0.34	5			

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
Prof Soldier	0.87	6	Health Care	0.43	5			
Education	0.97	6	Officials	0.46	5			
			House-wife	0.53	5			
			Educa-tion	0.60	5			
			Shipping Ind	0.22	6			
			Benefit WAO	0.23	6			
			Benefit AWW	0.30	6			
			Benefit WWW	1.25	6			
			Benefit AOW	1.27	6			
			Pensio-ner	1.40	6			
			Benefit VUT	1.53	6			
			Emp Gov Sub	-0.35	7			
			Benefit WW	-0.11	7			
			Benefit ABW	-0.09	7			
			Harbour Ind	-0.02	7			

Table A10. Coarse-classification by country. Goods code.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
BEDDING	-2.02	1	CAR EQUIP	-1.49	1	CARPETS	-1.48	0
KNIT MACHINE	-2.02	1	TELEPHON E	-0.98	2	93	-0.92	0
LAWN MOWERS	-1.32	1	CARPETS	-0.57	2	FLOWERS	-0.51	0
JEWEL-LERY	-1.32	1	CARD APPLICA	-0.35	2	TOOLS	-0.22	0
CAMERAS	-1.20	1	CAR RADIOS	-0.26	3	GARDEN FURN	0.10	0
OTHERS	-1.03	1	MOPEDS	-0.19	3	97	0.13	0
BEDROOM FURN	-0.76	1	OTHERS	-0.10	3	MOBIL-FUNK	-1.05	1
CYCLES	-0.76	1	HIFI RADIO	-0.09	3	PHONE FAX	-0.85	1
MOPEDS	-0.64	1	MISC BROWN G	-0.03	3	CYCLES	-0.82	1
OTHR PHONES	-0.59	2	CYCLES	0.00	3	SATEL ANT	-0.52	2
CAR RADIOS	-0.53	2	MICROWA VE	0.04	3	CAR EQUIP	-0.47	2
MOTORCY CLES	-0.50	2	OTH INSTRUME	0.07	4	HIFI RADIO	-0.43	2
LIVING FURN	-0.47	2	ORGANS	0.11	4	VIDEO RECORD	-0.26	3
HIFI RADIO	-0.32	2	TV	0.17	4	TV/VID COM	-0.17	3
BLTIN KITCHN	-0.31	2	MISC ELECT	0.19	4	HIFI RADIO	-0.15	3
COLOUR TV	-0.10	3	TV/VID COM	0.19	4	CYCLES	-0.10	3
MICROWA VE	-0.09	3	VIDEO RECORD	0.28	4	OTHERS	-0.04	3
TV/VIDEO COM	-0.07	3	COOKER	0.28	4	TV	0.12	4
DINING FURN	-0.07	4	DINING FURN	0.33	4	MICROWA VE	0.12	4
CD PLAYER	-0.01	4	PHOTO EQUIP	0.37	4	LIVING FURN	0.15	4
VACUUM CLNR	0.05	4	REFRIGER ATOR	0.40	4	BEDRM FURN	0.20	5
HSEHLD GOODS	0.05	4	LIVING FURN	0.41	4	REFRIGER ATOR	0.25	5
COMPUT REQ	0.13	5	VIDEO CAMERA	0.41	4	COMPUTE RS	0.26	5

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
MONO TV	0.14	5	BEDRM FURN	0.43	4	KITCHEN FURN	0.27	5
VIDEO CAMERA	0.20	5	WASH MACH	0.49	5	HIFI RADIO	0.28	5
HEATER	0.22	5	COMPUTE RS	0.49	5	WASH MACH	0.30	5
DRIER	0.26	5	FURNITUR E	0.50	5	HSEHLD GOODS	0.37	5
VIDEO RECORD	0.29	5	96	0.50	5	FURNITUR E	0.41	5
WASHING MACH	0.31	6	COMPUT REQ	0.51	5	PHOTO EQUIP	1.48	5
CARD APPLICA	0.32	6	DRIER	0.58	5	COOKER	1.74	5
COOKER	0.33	6	VACUUM CLNR	0.65	5			
COMPUTE RS	0.36	6	DISHWASH ER	0.76	6			
REFRIGER ATOR	0.38	6	MISC OTHERS	0.77	6			
KITCHEN FURN	0.55	7	FREEZER	0.85	6			
DISHWAS HER	0.60	7	CODE 97	0.88	6			
LADIES CLOTHES	0.82	7	SUNBEDS	1.10	6			
FREEZER	0.83	7	COMPUT GAME	-0.50	7			
SUNBEDS	1.26	7	LADIES CLOTHES	-0.46	7			
OTHERS	-0.02	8	63	-0.46	7			
			HSEHLD GOODS	-0.41	7			
			COMP DISC	-0.39	7			
			OTHR PHONES	-0.39	7			
			MISC OTHERS	-0.32	7			
			CD PLAYER	-0.11	7			
			CAMPING EQUP	0.15	7			
			FAX	0.31	7			
			95	0.35	7			
			LEATHER WEAR	0.49	7			

Table A11. Coarse-classification by country. Payment date.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
14	-2.92	1	01	0.34	1	01	0.14	1
22	-3.83	2	08	0.38	1	08	0.08	1
01	-3.95	3	15	0.28	2	15	-0.05	2
15	-3.99	3	22	-0.01	3	21	-0.40	3
00	-4.21	4	00	-1.15	4	22	-0.21	3
08	-4.21	4	28	-1.68	4			

Table A12. Coarse-classification by country. Number of dependants.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
0	-0.13	0	0	-0.77	1	0	-0.19	0
1	0.16	1	1	-0.35	2	1	0.22	1
2	0.28	2	2+	0.16	2	2	0.51	2
3	0.25	2	Unknown	0.01	3	3	0.24	3
4	-0.24	3				4	0.12	4
5+	-0.41	4				5+	0.07	4

Table A13. Coarse-classification by country. Spouse's age. 5% groups.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
18	-0.76	1	18	-0.01	1	0	0.43	0
23	-0.12	2	23	0.12	2	0	0.43	0
27	-0.01	3	27	0.10	2	0	0.43	0
30	0.14	4	28	0.17	2	18	-0.09	1
33	0.11	4	30	0.32	2	20	0.02	2
36	0.45	5	31	0.38	3	23	0.12	3
38	0.30	5	32	0.47	3	29	0.45	4
41	0.67	5	34	0.47	3	39	0.66	4
44	0.32	5	35	0.50	3	45	0.77	4
47	0.91	6	37	0.55	3	54	0.80	6
53	1.01	6	38	0.47	3	No Spouse	-0.27	5
63	1.80	6	40	0.60	4	No Spouse	-0.27	5
No Spouse	-0.27	7	41	0.65	4	No Spouse	-0.27	5
No Spouse	-0.27	7	43	0.60	4	No Spouse	-0.27	5
No Spouse	-0.27	7	45	0.84	5	No Spouse	-0.27	5
No Spouse	-0.27	7	47	0.91	5	No Spouse	-0.27	5
No Spouse	-0.27	7	50	0.73	5	No Spouse	-0.27	5
No Spouse	-0.27	7	53	1.00	5	No Spouse	-0.27	5
No Spouse	-0.27	7	59	1.63	5	No Spouse	-0.27	5
No Spouse	-0.27	7	No Spouse	-0.44	6	No Spouse	-0.27	5

Table A14. Coarse-classification by country. Time at address. 5% groups.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
No Info	-0.50	1	0	-0.52	0	0	-0.62	1
With Parents	-0.54	1	1m	-0.37	0	2m	-0.62	1
0	-0.60	1	3m	-0.38	0	4m	-0.73	1
4m	-0.38	2	5m	-0.34	0	6m	-0.50	2
1y	-0.28	2	8m	-0.43	0	10m	-0.29	2
1y4m	-0.29	2	1y1m	-0.23	1	1y	-0.63	2
1y11m	-0.09	3	1y3m	-0.11	1	1y4m	-0.27	2
2y6m	-0.26	3	1y7m	-0.06	1	1y9m	-0.13	2
3y1m	-0.21	3	2y	-0.33	1	2y1m	-0.43	2
3y9m	-0.05	4	2y4m	0.05	1	2y7m	-0.10	3
4y7m	0.11	4	2y10m	-0.19	1	3y	-0.19	3
5y7m	0.20	5	3y4m	0.06	2	3y9m	0.16	3
6y10m	0.28	5	4y1m	0.20	2	4y6m	0.04	4
8y4m	0.55	6	4y11m	0.04	2	5y8m	0.16	4
9y11m	0.49	6	5y11m	0.19	2	7y6m	0.19	4
11y8m	0.51	6	7y2m	0.28	2	9y7m	0.40	5
14y5m	1.09	6	9y1m	0.41	3	12y3m	0.54	5
18y1m	0.98	6	11y1m	0.52	3	16y2m	0.74	5
22y1m	0.87	6	14y5m	0.67	3	22y2m	0.81	6
31y4m	1.50	6	19y8m	0.83	3	53y	1.29	6

Table A15. Coarse-classification by country. Time on job. 5% groups.

BELGIUM			NETHERLANDS			GERMANY		
CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP	CHAR-C	WOE	GROUP
0	-0.78	1	0	-0.45	1	0	0.86	0
5m	-0.55	2	3m	-0.55	1	1m	-0.70	1
10m	-0.58	2	5m	-0.47	1	3m	-0.74	1
1y4m	-0.42	2	8m	-0.32	2	5m	-0.65	1
2y	-0.29	3	11m	-0.31	2	7m	-0.63	2
2y7m	-0.28	3	1y3m	-0.26	2	11m	-0.51	2
3y4m	-0.17	3	1y7m	-0.23	2	1y4m	-0.40	2
4y2m	-0.29	3	2y1m	-0.26	2	2y	-0.37	2
5y5m	0.03	3	2y8m	-0.11	3	2y5m	-0.44	2
7y2m	0.39	4	3y3m	-0.09	3	3y	-0.42	3
8y9m	0.27	4	4y1m	0.06	3	3y8m	-0.32	3
10y5m	0.48	5	5y2m	0.09	3	4y8m	-0.08	3
13y2m	0.67	5	6y7m	0.23	4	6y	0.29	4
17y7m	0.88	6	8y2m	0.30	4	7y7m	0.48	4
21y3m	1.13	6	9y1m1	0.32	4	9y5m	0.72	5
27y10m	0.94	6	12y6m	0.45	4	13y6m	0.65	5
32y	1.09	6	16y9m	0.60	5	22y5m	0.86	5
Retired	0.83	6	21y	0.72	5	Agency	0.83	6
House-wife	0.11	7	31y	-0.13	5	Retired	0.12	6
Allow-ance	-0.06	7	Retired	0.74	5	House-wife	0.01	6

Table A16. Coarse-classification by country. Goods price. 5% groups.

BELGIUM			NETHERLANDS			GERMANY		
BF	WOE	GROUP	NLG	WOE	GROUP	DM	WOE	GROUP
0	0.35	0	0	0.01	0	0	-0.10	0
6000	0.02	1	100	-1.00	1	450	-0.51	1
8990	0.03	1	415	-0.60	1	557	-0.42	2
9998	-0.07	1	565	-0.25	2	649	-0.35	2
11998	-0.06	1	700	-0.23	2	700	-0.25	3
13500	0.12	2	837	-0.02	2	799	-0.28	3
14998	0.16	2	1000	0.04	3	899	-0.19	3
16995	0.24	2	1154	0.19	3	999	-0.07	4
18000	0.01	2	1350	0.03	3	1099	-0.10	4
19990	0.09	2	1568	0.15	3	1214	0.00	4
20385	0.50	3	1858	0.21	3	1382	0.12	4
22995	0.24	3	2146	0.05	3	1499	0.13	4
24998	-0.03	4	2500	0.13	3	1638	0.09	4
27588	-0.07	4	2800	0.39	4	1800	0.19	5
30000	0.21	5	3020	0.42	4	1999	0.26	5
34792	-0.27	6	3229	0.43	4	2139	0.17	5
39974	-0.35	6	3631	0.49	4	2399	0.33	5
40000	-0.55	6	3996	0.50	4	2761	0.21	5
50000	-0.03	7	4596	0.23	4	3459	0.14	5
60000	0.05	7	5362	0.33	4	90000	0.22	5

A17. Binary variable coding for Belgium

```
        phone1=0; phone2=0;          /*Home phone given*/
    select (phone_g);
        when (1)      phone1=1;
        when (2)      phone2=1;
    otherwise;
    end;

        rstat1=0; rstat2=0; rstat3=0;           /*Residential status*/
select (rstat);
        when ('1')      rstat1=1;
        when ('2')      rstat2=1;
        when ('3')      rstat3=1;
otherwise;
    end;

        mstat1=0; mstat2=0;          /*Marital status*/
select (mstat);
        when ('1','2') mstat1=1;
        when ('3')      mstat3=1;
otherwise;
    end;

        occup1=0; occup2=0; occup3=0; occup4=0; occup5=0;   /*Occupation*/
    select (occup_g);
        when (1)      occup1=1;
        when (2)      occup2=1;
        when (3)      occup3=1;
        when (4)      occup4=1;
        when (5)      occup5=1;
    otherwise;
    end;

        age1=0; age2=0; age3=0; age4=0; age5=0;          /*Applicant's age*/
    select (age_g);
        when (1)      age1=1;
        when (2)      age2=1;
        when (3)      age3=1;
        when (4)      age4=1;
        when (5)      age5=1;
    otherwise;
    end;
```

```
        bus11=0; bus12=0; bus13=0; bus14=0; bus15=0; bus16=0;
    select (bus1_g);                  /*Business type*/
        when (1)      bus11=1;
        when (2)      bus12=1;
        when (3)      bus13=1;
        when (4)      bus14=1;
        when (5)      bus15=1;
        when (6)      bus16=1;
    otherwise;
    end;

        kids10=0; kids11=0; kids12=0; kids13=0;  /*Number of dependants*/
    select (kids1_g);
        when (0)      kids10=1;
        when (1)      kids11=1;
        when (2)      kids12=1;
        when (3)      kids13=1;
    otherwise;
    end;

        gdsc1=0; gdsc2=0; gdsc3=0; gdsc4=0; gdsc5=0; gdsc6=0;
select (gdsc_g);                      /*Goods code*/
        when (1)      gdsc1=1;
        when (2)      gdsc2=1;
        when (3)      gdsc3=1;
        when (4)      gdsc4=1;
        when (5)      gdsc5=1;
        when (6)      gdsc6=1;
otherwise;
    end;

gp0=0; gp1=0; gp2=0; gp3=0; gp4=0; gp5=0; gp6=0;
    select (gp_g);                    /*Goods price*/
        when (0)      gp0=1;
        when (1)      gp1=1;
        when (2)      gp2=1;
        when (3)      gp3=1;
        when (4)      gp4=1;
        when (5)      gp5=1;
        when (6)      gp6=1;
    otherwise;
    end;
```

```
     spage1=0; spage2=0; spage3=0; spage4=0; spage5=0; spage6=0;
select (spage_g);                          /*Spouse's age*/
     when (1)     spage1=1;
     when (2)     spage2=1;
     when (3)     spage3=1;
     when (4)     spage4=1;
     when (5)     spage5=1;
     when (6)     spage6=1;
otherwise;
end;

     cardi0=0;                       /*Card insurance*/
select (cardi);
     when (0)     cardi0=1;
otherwise;
end;

     credi0=0;                       /*Credit insurance*/
select (credi);
     when (0)     credi0=1;
otherwise;
end;

     pdate8=0; pdate14=0; pdate15=0;            /*Payment date*/
select (ppdate);
     when ('00','08')     pdate8=1;
     when ('14')          pdate14=1;
     when ('01','15')     pdate15=1;
otherwise;
end;

     tad1=0; tad2=0; tad3=0; tad4=0; tad5=0;
select (tad_g);                                  /*Time at address*/
     when (1)     tad1=1;
     when (2)     tad2=1;
     when (3)     tad3=1;
     when (4)     tad4=1;
     when (5)     tad5=1;
otherwise;
end;
```

```
          toj1=0; toj2=0; toj3=0; toj4=0; toj5=0; toj6=0;
     select (toj_g);                              /*Time in employment*/
          when (1)      toj1=1;
          when (2)      toj2=1;
          when (3)      toj3=1;
          when (4)      toj4=1;
          when (5)      toj5=1;
          when (6)      toj6=1;
otherwise;
     end;

ephone0=0;          /*Employer's phone given*/
     select (ephone);
          when (1)            ephone0=1;
     otherwise;
     end;
```

Additional variables for 'full information' model

```
agreem1=0; agreem2=0;       /*Type of agreement*/
     select (agreem);
          when ('01')                 agreem1=1;
          when ('02','47','60')       agreem2=1;
     otherwise;  /* other */
     end;

ident1=0; ident2=0;         /*Type of ID */
     select (ident);
          when ('01','02')            ident1=1;
          when ('03')                 ident2=1;
     otherwise; /*other */
     end;

lang1=0; lang2=0;           /*Language*/
     select (lang);
          when ('F')          lang1=1;
          when ('N')          lang2=1;
     otherwise; /*other*/
     end;

premp0=0;                   /*Previous employment*/
     select (premp);
          when (0)           premp0=1;
     otherwise;
     end;
```

```
tcont0=0;                       /*Total contracts*/
      select (tcont);
            when (0)          tcont0=1;
      otherwise;
      end;

under1=0; under2=0;                  /*Underwriter*/
      select (under);
            when ('401')                 under1=1;
            when ('009','013')           under2=1;
      otherwise;
      end;

taddr1=0; taddr2=0;                  /*Total addresses given*/
      select (taddr);
            when (1)                     taddr1=1;
            when (2)                     taddr2=1;
      otherwise; /*3 */
      end;

ininst0=0; ininst1=0;  ininst2=0;            /*Initial installment*/
      select;
            when (ininst=0)                      ininst0=1;
            when (ininst=1000)                   ininst1=1;
            when (1500<=ininst<1900)             ininst2=1;
      otherwise;
      end;

t_fb0=0;                                     /*Time first bureau*/
      select;
            when (t_fb=0)             t_fb0=1;

      otherwise;
      end;

card1=0; card2=0;                    /*Credit card type*/
        select (card);
            when (1)              card1=1;
            when (2)              card2=1;
      otherwise; /*no card*/
      end;
```

```
tbank1=0; tbank2=0; tbank3=0; tbank4=0; tbank5=0; tbank6=0;     /*Time at bank*/
    select;
        when (0<=tbank<200)           tbank1=1;
        when (200<=tbank<300)         tbank2=1;
        when (300<=tbank<500)         tbank3=1;
        when (500<=tbank<700)         tbank4=1;
        when (700<=tbank<1000)        tbank5=1;
        when (1000<=tbank<1600)       tbank6=1;
    otherwise;  /* 16+ */
    end;

spocc1=0; spocc2=0; spocc3=0;       /*Spouse's occupation*/
    select (spocc);
        when ('5','9')          spocc1=1;
        when ('1','8')          spocc2=1;
        when ('2','3','4')      spocc3=1;
    otherwise; /*no spouse*/
    end;

retail1=0; retail2=0; retail3=0;  retail4=0;      /*Retailer*/
    select (retail);
        when (70999)                                   retail1=1;
        when (50199,100023,240199)                     retail2=1;
        when (0)                                       retail3=1;
        when (60199,130199,200199,210598,210599)       retail4=1;
    otherwise; /*other*/
    end;
```

A18. Binary variable coding for the Netherlands

```
ephone0=0;            /*Employer's phone given*/
    select (ephone);
        when (2)      ephone0=1;  /*telephone not given */
    otherwise;
    end;

        phone1=0;  /*Home phone given*/
    select (phone);
        when (1,2)      phone1=1;  /*given, secret */
    otherwise;                     /*mobile, not given, missing */
    end;
```

```
          rstat1=0; rstat2=0; rstat3=0; rstat4=0;       /*Residential status*/
     select (rstat);
          when ('01')                   rstat3=1;   /*rented house */
          when ('02')                   rstat4=1;   /*home owner*/
          when ('03','99','06')         rstat2=1;   /*rented flat, no info */
          when ('04')                   rstat1=1;   /*living with parents */
     otherwise;                        /*caravan, rented room, missing */
     end;

          mstat1=0; mstat2=0; mstat3=0;      /*Marital status*/
     select (mstat);
          when ('1')          mstat3=1;      /*married */
          when ('5','6')      mstat2=1;      /*widowed, living together reg*/
          when ('3','4')      mstat1=1;      /*divorced, living together */
     otherwise;                          /*single, missing */
     end;

     age1=0; age2=0; age3=0; age4=0; age5=0;  age6=0;
 select;                          /*Applicant's age*/

     when (19<age<23)      age1=1;
     when (22<age<26)      age2=1;
     when (25<age<30)      age3=1;
     when (29<age<33)      age4=1;
     when (32<age<43)      age5=1;
     when (42<age<54)      age6=1;
otherwise; /* <53+ */
end;

          btype1=0;  btype2=0; btype3=0; btype4=0; btype5=0; btype6=0;
     select (btype);              /*Business type*/
          when ('63')                             btype1=1;      /*agency*/
          when ('05','09','12','51','62','55')    btype2=1;      /*military prof, catering,
                                        farming, benefit AAW, RWW, student+job*/
          when ('10','01','17','03','14')         btype3=1;      /*industry, building trd,
                                        shopwork, road trp, manual wrk */
          when ('99','18','20','08','19','11')    btype4=1;      /*unknown,
                    businessman, cleaning, computer ind, bank/ins, service ind */
          when ('16','07','66','06','02','04')    btype5=1;
                   /*airman, education, housewife, pub health, gov off, pub trans*/
          when ('53','60','64')                     btype6=1;
                    /*benefit AOW, VUT, pensioner*/
          otherwise; /*military serv, emp gov sub, dock wrk, seamen, benefit WWW,
                      ROA, indep means, benefit WW, ABW, WAO, AWW*/
end;
```

```
     gcode1=0; gcode2=0; gcode3=0; gcode4=0; gcode5=0; gcode6=0;
  select (gcode);                                /*Goods code*/
     when ('64')                              gcode1=1; /*car equip*/
     when ('00','44','30')                    gcode2=1; /* card app, carpets,
                                                        telephone*/
     when ('50','19','03','09','51','16','06')   gcode3=1;
              /*mopeds,misc,hifi,misc brown g,cycles,microwave,car hifi*/
     when ('20','10','42','41','15','40','21')   gcode4=1;
                   /*photo equip,fridge,furniture,cooker,videocamera*/
     when ('11','27','28','49','95','12','17')   gcode5=1;
                   /*wash,comp,comp equip,misc furn,95,drier,vac*/
     when ('90','14','13','97','23')          gcode6=1;
                   /*pers loan, dishwasher,freezer,misc97, sun lamps */
  otherwise;       /*small categories, misc music,organs,misc elct, tv, video*/
  end;

  spage1=0; spage2=0; spage3=0; spage4=0;          /*Spouse's age*/
  select;
  when (17<spage<23)    spage1=1;
  when (22<spage<31)    spage2=1;
  when (30<spage<45)    spage3=1;
  when (44<spage<999)   spage4=1;
otherwise;   /* 999 */
end;

     cardi0=0;                    /*Card insurance*/
  select (cardi);
     when (0)      cardi0=1;  /*no insurance */
  otherwise;
  end;                       /*insurance */

     credi0=0; credi1=0;          /*Credit insurance*/
  select;
     when (credi=0)      credi0=1;    /* no insurance */
     when (credi=8)      credi1=1;    /* credi 8 */

  otherwise;                        /*insurance, other */
  end;

     pdate1=0; pdate2=0; pdate3=0;     /*Payment date*/
  select (pdate);
     when ('01')  pdate1=1;
     when ('08')  pdate2=1;
     when ('15','22')  pdate3=1;
  otherwise;                    /* 00, 28 */
  end;
```

```
        tad1=0; tad2=0; tad3=0;              /*Time at address*/
    select;
    when (6<tad<207)         tad1=1;
    when (206<tad<900)       tad2=1;
    when (899<tad)           tad3=1;
otherwise;   /*tad=0 and <7 */
    end;

kids0=0; kids1=0;                            /*Number of dependants*/
    select;
        when (kids=0)        kids0=1;        /*no kids */
        when (0<kids<99)  kids1=1;           /*kids */
    otherwise;                               /*no info */
    end;

        gpr1=0; gpr2=0; gpr3=0; gpr4=0;
select;                                      /*Goods price*/
    when (0<gpr<500)              gpr1=1;
    when (499<gpr<1000)           gpr2=1;
    when (999<gpr<3000)           gpr3=1;
    when (2999<gpr)               gpr4=1;
otherwise;  /*gpr=0 */
    end;

        occup1=0; occup2=0; occup3=0;  /*Occupation*/
    select (occup);

        when ('01')              occup3=1;                /*employed */
        when ('02')              occup1=1;                /*self-emp */
        when ('05','06','07','08')   occup2=1;            /*agency */
    otherwise;                          /*part-time, bw, hw, benefit */
    end;

        toj1=0; toj2=0; toj3=0; toj4=0; toj5=0;
    select;                                  /*Time in employment*/
    when (toj<7)             toj1=1;
    when (6<toj<207)         toj2=1;
    when (206<toj<600)       toj3=1;
    when (599<toj<1500)      toj4=1;
otherwise;                   /*1499<toj<9990, retired*/
end;
```

```
/*Interactions between Goods Code and Payment Date*/
if gcode_g=1 and pdate_g=1 then gcpd11=1; else gcpd11=0;
if gcode_g=1 and pdate_g=2 then gcpd12=1; else gcpd12=0;
if gcode_g=1 and pdate_g=3 then gcpd13=1; else gcpd13=0;
if (gcode_g=1 and pdate_g=4) or (gcode_g=3 and pdate_g=4)
or (gcode_g=4 and pdate_g=4) or (gcode_g=5 and pdate_g=4) then gcpd14=1; else
gcpd14=0;

if (gcode_g=2 and pdate_g=1) or (gcode_g=2 and pdate_g=3) then gcpd21=1; else
gcpd21=0;
if gcode_g=2 and pdate_g=2 then gcpd22=1; else gcpd22=0;

if gcode_g=2 and pdate_g=4 then gcpd24=1; else gcpd24=0;

if gcode_g=3 and pdate_g=1 then gcpd31=1; else gcpd31=0;
if gcode_g=3 and pdate_g=2 then gcpd32=1; else gcpd32=0;
if gcode_g=3 and pdate_g=3 then gcpd33=1; else gcpd33=0;

if gcode_g=4 and pdate_g=1 then gcpd41=1; else gcpd41=0;
if gcode_g=4 and pdate_g=2 then gcpd42=1; else gcpd42=0;
if gcode_g=4 and pdate_g=3 then gcpd43=1; else gcpd43=0;


if gcode_g=5 and pdate_g=1 then gcpd51=1; else gcpd51=0;
if (gcode_g=5 and pdate_g=2) or (gcode_g=6 and pdate_g=2) then gcpd52=1; else
gcpd52=0;
if gcode_g=5 and pdate_g=3 then gcpd53=1; else gcpd53=0;


if gcode_g=6 and pdate_g=1 then gcpd61=1; else gcpd61=0;
if gcode_g=6 and pdate_g=3 then gcpd63=1; else gcpd63=0;
if gcode_g=6 and pdate_g=4 then gcpd64=1; else gcpd64=0;

if gcode_g=7 and pdate_g=1 then gcpd71=1; else gcpd71=0;
if gcode_g=7 and pdate_g=2 then gcpd72=1; else gcpd72=0;
if gcode_g=7 and pdate_g=3 then gcpd73=1; else gcpd73=0;
```

Additional variables for 'full information' model

```
loan1=0; loan2=0; loan3=0; loan4=0;          /*Loan amount*/
select ;
    when (0<loan<250)            loan1=1;
    when (250<=loan<500)         loan2=1;
    when (500<=loan<750)         loan3=1;
    when (750<=loan<1500)        loan4=1;
    when (loan=999999)           loan4=1;
    otherwise; end;
```

```
instal0=0; instal1=0;            /*Instalment paid */
select ;
    when (instal=0)                  instal0=1;
    when (0<instal<3000)             instal1=1;
    otherwise;
    end;

negbu1=0; negbu2=0; negbu3=0;              /*Negative bureau */
select (negbu) ;
    when ('UN')            negbu1=1;
    when ('02')            negbu2=1;
    when ('03')            negbu3=1;
    otherwise;
    end;


paidft0=0; paidft1=0; paidft2=0;       /*Number of paid fixed term accounts */
select (paidft) ;
    when (0,99999)         paidft0=1;
    when (1)               paidft1=1;
    when (2,3,4)           paidft2=1;
    otherwise;
    end;


livft0=0; livft1=0;      /*Number of live fixed term accounts*/
select (livft) ;
    when (0)               livft0=1;
    when (.,99999)         livft1=1;
    otherwise;
    end;


livecr0=0; livecr1=0; livecr2=0;       /*Number of live revolving accounts */
select (livecr) ;
    when (0)               livecr0=1;
    when (1)               livecr1=1;
    when (2,99999)         livecr2=1;
    otherwise;
    end;
```

```
lasta0=0; lasta1=0; lasta2=0; lasta3=0; lasta4=0; lasta5=0; lasta6=0;
select ;                          /*Time since last A account*/
     when (lasta=0)                      lasta0=1;
     when (0<lasta<4)                    lasta1=1;
     when (3<lasta<9)                    lasta2=1;
     when (8<lasta<100)                  lasta3=1;
     when (100<lasta<200)                lasta4=1;
     when (200<lasta<600)                lasta5=1;
     when (lasta=99999)                  lasta6=1;
     otherwise;
     end;

new0=0; new1=0;       /*Number of new accounts in last 6 months*/
select (new) ;
     when (99999)              new1=1;
     when (0)                  new0=1;
     otherwise;
     end;


f_buro1=0; f_buro2=0; f_buro3=0; f_buro4=0; f_buro5=0;
select ;             /*Time first bureau registration */

     when (f_buro=99999)                 f_buro1=1;
     when (0<=f_buro<8)                  f_buro2=1;
     when (7<f_buro<100)                 f_buro3=1;
     when (100<=f_buro<204)              f_buro4=1;
     when (203<f_buro<400)               f_buro5=1;
     otherwise;
     end;

lastap0=0; lastap1=0;  lastap2=0; lastap3=0;
select ;                    /* Time since last A+ account*/
     when (lastap=9999)                  lastap0=1;
     when (lastap=99999)                 lastap1=1;
     when (0<=lastap<109)                lastap2=1;
     when (109<=lastap<612)              lastap3=1;
     otherwise;
     end;

last_a0=0; last_a1=0;             /*Last A account*/
select ;
     when (last_a=9999)                  last_a0=1;
     when (last_a=99999)                 last_a1=1;
     otherwise;
     end;
```

```
cl_b0=0; cl_b1=0;              /*Total closed bureau*/
select (cl_b);
     when (0)                  cl_b0=1;
     when (99999)              cl_b1=1;
     otherwise;
     end;

liveb0=0; liveb1=0;            /*Total live bureau*/
select (liveb);
     when (0)                  liveb0=1;
     when (99999)              liveb1=1;
     otherwise;
     end;

aplus0=0; aplus1=0;            /*Total A+ accounts */

select (aplus);
     when (0)                  aplus0=1;
     when (99999)              aplus1=1;
     otherwise;
     end;

a_acc0=0; a_acc1=0;            /*Total A accounts */
select (a_acc);
     when (0)                  a_acc0=1;
     when (99999)              a_acc1=1;
     otherwise;
     end;

depos0=0;                      /*Deposit */
     select;
     when (depos=0)          depos0=1;
     otherwise;
     end;

deal1=0; deal2=0; deal3=0;     /*Dealer */
     select;
     when (deal=16)            deal1=1;
     when (20<deal<24)         deal1=1;
     when (48<deal<56)         deal1=1;
     when (0<=deal<5)          deal2=1;
     when (9<deal<13)          deal2=1;
     when (29<=deal<35)        deal2=1;
     when (42<deal<49)         deal2=1;
     when (90<=deal<94)        deal3=1;
     otherwise;
     end;
```

```
sect1=0;               /*Section */
     select;
     when (200<sect<400)    sect1=1;
     otherwise;
     end;

bacc1=0; bacc2=0;      /*Type of bank account */
     select (bankacc);
     when (1)          bacc1=1;
     when (2)          bacc2=1;
     otherwise;
     end;

bankc1=0; bankc2=0; bankc3=0; bankc4=0;          /*Bank code */
     select (bankc);
     when (06)         bankc1=1;
     when (05)         bankc2=1;
     when (04)         bankc3=1;
     when (01,02,09) bankc4=1;
     otherwise;
     end;

nation1=0;             /*Nationality */
     select (nation);
     when (01)         nation1=1;
     otherwise;
     end;

agr1=0; agr2=0; agr3=0; agr4=0; agr5=0;     /*Type of agreement */
     select (agr);
     when (33,36)                           agr1=1;
     when (11,13,17,14,18,12,15,91)         agr2=1;
     when (01,02,03,04,80,81,82)            agr3=1;
     when (21,23,24,26,27,28,29)            agr4=1;
     when (09,08)                           agr5=1;
     otherwise;
     end;

spocc1=0; spocc2=0; spocc3=0; spocc4=0;  /*Spouse's occupation */
     select (spocc);
     when ('NS')              spocc1=1;
     when ('20')              spocc2=1;
     when ('08','07')         spocc3=1;
     when ('01','02')         spocc4=1;
     otherwise;
end;
```

A19. Binary variable coding for Germany

```
        ephone0=0;          /*Employer's phone given*/
    select (ephone);
        when (1)            ephone0=1;
    otherwise;
    end;

        rstat1=0; rstat2=0; rstat3=0;         /*Residential status*/
    select (rstat);
         when('1')           rstat1=1;
         when('2','6')       rstat2=1;
         when('3')           rstat3=1;
         otherwise;
end;
        mstat1=0; mstat2=0;                   /*Marital status*/
    select (mstat);
         when('1','4')       mstat1=1;
         when('2')           mstat2=1;
         otherwise;
end;

         phone1=0;                            /*Home phone given*/
    select (phone);
         when('0')           phone1=1;
         otherwise;
end;

        kids0=0; kids1=0; kids2=0; kids3=0;         /*Number of dependants */
    select (kids);
         when(0)    kids0=1;
         when(1)    kids1=1;
         when(2)    kids2=1;
         when(3)    kids3=1;
         otherwise;
end;

        btype1=0; btype2=0; btype3=0; btype4=0; btype5=0; btype6=0;
    select (btype);                          /*Business type*/
         when('06')                 btype1=1;
         when('12','05','07','08')  btype2=1;
         when('04','14','10')       btype3=1;
         when('09','13')            btype4=1;
         when('01','02')            btype5=1;
         when('03','11')            btype6=1;
         otherwise;
end;
        occ1=0; occ2=0; occ3=0; occ4=0;             /*Occupation */
```

```
        select (occ);
              when('06','10')              occ1=1;
              when('09','07','08')         occ2=1;
              when('02')                   occ3=1;
              when('01','04','05')         occ4=1;
              otherwise;
end;

             cardi1=0;                            /*Card Insurance */
        select (cardi);
              when('1')                    cardi1=1;
              otherwise;
end;

        age1=0; age2=0; age3=0; age4=0; age5=0; age6=0;     /*Applicant's age */
        select ;
              when(17<age<20)         age1=1;
              when(19<age<22)         age2=1;
              when(21<age<27)         age3=1;
              when(26<age<32)         age4=1;
              when(31<age<39)         age5=1;
              when(39<age<60)         age6=1;
              otherwise;
end;

         spage0=0; spage1=0; spage2=0; spage3=0; spage4=0; spage5=0;
              select ;                             /*Spouse's age */
              when(spage=0)              spage0=1;
              when(17<spage<20)          spage1=1;
              when(19<spage<23)          spage2=1;
              when(22<spage<29)          spage3=1;
              when(28<spage<36)          spage4=1;
              when(35<spage<40)          spage4=1;
              when(39<spage<45)          spage4=1;
              when(44<spage<49)          spage4=1;
              when(48<spage<55)          spage4=1;
              when(spage_g=5)            spage5=1;
              otherwise;
end;
```

```
    tad1=0; tad2=0; tad3=0; tad4=0; tad5=0;          /*Time at address */
          select;
          when(tad<=5)              tad1=1;
          when(5<tad<207)           tad2=1;
          when(206<tad<404)         tad3=1;
          when(403<tad<=811)        tad4=1;
          when(811<tad<=2111)       tad5=1;
          otherwise;
end;

       toj1=0; toj2=0; toj3=0; toj4=0; toj6=0;       /*Time in employment */
         select;
          when(0<=toj<=6)           toj1=1;
          when(6<toj<=210)          toj2=1;
          when(210<toj<=410)        toj3=1;
          when(410<toj<=906)        toj4=1;
          when(toj=9995)            toj6=1;
          when(toj=9997)            toj6=1;
          when(toj=9998)            toj6=1;
          otherwise;
end;

    gprice0=0; gprice1=0; gprice2=0; gprice3=0; gprice4=0;
    select;                           /*Goods price */
          when(gprice=0)                    gprice0=1;
          when(0<gprice<=450)               gprice1=1;
          when(450<gprice<=649)             gprice2=1;
          when(649<gprice<=899)             gprice3=1;
          when(899<gprice<=1638)            gprice4=1;
          otherwise;
end;

    gcode1=0; gcode2=0; gcode3=0; gcode4=0; gcode5=0;
select   (gcode);                     /*Goods code */
          when('24','25','43')                                       gcode1=1;
          when('23','40','05')                                       gcode2=1;
          when('02','03','08','41','42','90')                        code3=1;
          when('01','11','97','34')                                  gcode4=1;
          when('04','33','12','20','38','06','13','36','19','30','26','14')  gcode5=1;
          otherwise;
end;
```

```
    pday1=0; pday2=0;                /*Payment date */
    select (pday);
          when('01','08','21')              pday1=1;
          when('15')                        pday2=1;
          otherwise;
end;

    credi1=0;                        /*Credit insurance */
      select (credi);
          when('1')           credi1=1;
otherwise;
end;
```

Additional variables for 'full information' model

```
ininst1=0; ininst2=0; ininst3=0; ininst4=0; ininst5=0;  ininst6=0;
    select;                     /*Initial instalment */

          when(0<=ininst<35)                ininst1=1;
          when(35<=ininst<50)               ininst2=1;
          when(ininst=50)                   ininst3=1;
          when(51<=ininst<100)              ininst4=1;
          when(100<=ininst<200)             ininst5=1;
          when(200<=ininst<300)             ininst6=1;
          otherwise;
end;

l_cr0=0; l_cr1=0; l_cr2=0; l_cr3=0;            /* Total amount of live credit */
    select;
          when(l_cr=0)                      l_cr0=1;
          when(1<=l_cr<50)                  l_cr1=1;
          when(50<=l_cr<70)                 l_cr2=1;
          when(70<=l_cr<200)                l_cr3=1;
          otherwise;
end;

loan0=0; loan1=0; loan2=0; loan3=0;  loan4=0;      /*Loan amount */
    select;
          when(loan=0)                      loan0=1;
          when(1<=loan<500)                 loan1=1;
          when(500<=loan<700)               loan2=1;
          when(700<=loan<900)               loan3=1;
          when(900<=loan<1200)              loan4=1;
          otherwise;
end;
```

```
lam_300=0; lam_301=0; lam_302=0; lam_303=0; lam_304=0;
    select;                    /*Loan amount 30*/
        when(lam_30=0)               lam_300=1;
        when(1<=lam_30<18)           lam_301=1;
        when(18<=lam_30<30)          lam_302=1;
        when(30<=lam_30<50)          lam_303=1;
        when(50<=lam_30<100)         lam_304=1;
        when(100<=lam_30<150)        lam_303=1;
        otherwise;
end;

depos0=0; depos1=0; depos2=0; depos3=0;          /*Deposit*/
    select;
        when(depos=0)                depos0=1;
        when(1<=depos<100)           depos1=1;
        when(100<=depos<200)         depos2=1;
        when(200<=depos<500)         depos3=1;
        otherwise;
end;

tbank1=0; tbank2=0; tbank3=0; tbank4=0; tbank5=0;          /*Time at bank*/
    select;
        when(0<=tbank<6)             tbank1=1;
        when(6<=tbank<200)           tbank2=1;
        when(200<=tbank<300)         tbank3=1;
        when(300<=tbank<500)         tbank4=1;
        when(600<=tbank<900)         tbank5=1;
        otherwise;
end;

timepa0=0; timepa1=0; timepa2=0;             /*Time at previous address */
    select;
        when(timepa=.)               timepa0=1;
        when(1<=timepa<200)          timepa1=1;
        when(200<=timepa<500)        timepa2=1;
        otherwise;
end;

timepj0=0; timepj1=0; timepj2=0; timepj3=0;          /*Time in previous job */
    select;
        when(timepj=.)               timepj0=1;
        when(1<=timepj<100)          timepj1=1;
        when(100<=timepj<200)        timepj2=1;
        when(200<=timepj<500)        timepj3=1;
        otherwise;
end;
```

```
btim1=0; btim2=0; btim3=0; btim4=0; btim5=0;      /*Time at bureau */
     select;
          when(btim<5)                    btim1=1;
          when(5<=btim<200)               btim2=1;
          when(200<=btim<400)             btim3=1;
          when(9800<=btim<9905)           btim4=1;
          when(btim=9998)                 btim5=1;
          otherwise;
end;

t1b0=0; t1b1=0; t1b2=0; t1b3=0; t1b4=0; t1b5=0;           /*Time at bureau 1*/
     select;
          when(t1b=0)                     t1b0=1;
          when(1<=t1b<8)                  t1b1=1;
          when(8<=t1b<200)                t1b2=1;
          when(9800<=t1b<9905)            t1b3=1;
          when(9905<=t1b<9998)            t1b4=1;
          when(t1b=9998)                  t1b3=1;
          when(t1b=9999)                  t1b4=1;
          when(t1b=.)                     t1b5=1;
          otherwise;
end;

t2b1=0; t2b2=0; t2b3=0; t2b4=0; t2b5=0;           /*Time at bureau 2*/
     select;

          when(t2b<6)                     t2b1=1;
          when(6<=t2b<200)                t2b2=1;
          when(200<=t2b<300)              t2b3=1;
          when(300<=t2b<800)              t2b4=1;
          when(t2b=9998)                  t2b5=1;
          otherwise;
end;


bscod1=0; bscod2=0; bscod3=0; bscod4=0;          /*Bank sort code */
     select (bscod);
          when(1)                 bscod1=1;
          when(0,2,3)             bscod2=1;
          when(5)                 bscod3=1;
          when(6,9)               bscod4=1;

          otherwise;
end;
```

```
s_dl1=0; s_dl2=0; s_dl3=0; s_dl4=0;          /*Section, dealer */
    select (s_dl);
        when(12,20,26)                 s_dl1=1;
        when(10,11,15,3,14,16)         s_dl2=1;
        when(0,9)                      s_dl3=1;
        when(6,8)                      s_dl4=1;
        otherwise;
end;

thinst1=0; thinst2=0;                   /*Theoretical instalment */
     select;
         when (thinst=. )              thinst2=1;
         when (0<=thinst<2)            thinst1=1;
         when (thinst=2)               thinst2=1;
         when (9<=thinst<16)           thinst2=1;
     otherwise;
    end;

pdep0=0; pdep1=0; pdep2=0; pdep3=0; pdep4=0;    /*Percent deposit */
     select;
         when (0<pdep<11)              pdep1=1;
         when (pdep=0)                 pdep0=1;
         when (11<=pdep<21)            pdep2=1;
         when (21<=pdep<30)            pdep3=1;
         when (30<=pdep<43)            pdep4=1;
         when (pdep=97)                pdep2=1;
     otherwise;
    end;

l_init1=0; l_init2=0; l_init3=0; l_init4=0;     /*Loan to instalment */
     select;
         when (0<=l_init<3)            l_init1=1;
         when (3<=l_init<11)           l_init2=1;
         when (11<=l_init<16)          l_init3=1;
         when (16<=l_init<30)          l_init4=1;
     otherwise;
    end;

btinc0=0; btinc1=0; btinc2=0;           /*Bureau total increment */
     select (btinc) ;
         when (.)           btinc0=1;
         when (1,2,3,4)     btinc1=1;
         when (99)          btinc2=1;
     otherwise;
    end;
```

```
under1=0; under2=0;            /*Underwriter */
      select (under) ;
          when ('011', '012', '019')          under1=1;
          when ('001', ' ')                   under2=1;
      otherwise;
    end;


agreem1=0;                     /*Type of agreement */
      select (agreem) ;
          when ('01')          agreem1=1;
otherwise;
    end;

ident1=0;
      select (ident) ;         /*Type of ID*/
          when ('1')          ident1=1;
      otherwise;
    end;

ttlad1=0; ttlad2=0;            /*Total addresses given */
      select (ttlad) ;
          when (1)          ttlad1=1;
          when (2)          ttlad2=1;
      otherwise;
    end;

spouse0=0;                     /*Spouse marker */
      select (spouse) ;
          when ('0')          spouse0=1;
      otherwise;
    end;

crcard1=0; crcard2=0;          /*Credit card type */
      select (crcard) ;
          when (1)          crcard1=1;
          when (2)          crcard2=1;
      otherwise;
    end;

cardh1=0;                      /*2nd card holder */
      select (cardh2) ;
          when ('1')          cardh1=1;
      otherwise;
    end;
```

```
e_w1=0; e_w2=0; e_w3=0;   /*East/West Germany */
      select (e_w);
            when ('EA')          e_w1=1;
            when ('BE')          e_w2=1;
            when ('WE')          e_w3=1;
      otherwise;
     end;

typeb1=0;                 /*Type of bank account */
      select (typeb) ;
            when ('1')          typeb1=1;
      otherwise;
     end;

cardst0=0;                /*Card status */
      select (cardst) ;
            when ('0')          cardst0=1;
      otherwise;
     end;

prins0=0;                 /*Product insurance */
      select (prins) ;
            when ('0')          prins0=1;
      otherwise;
     end;

b_cred1=0; b_cred2=0;     /*Worst bureau credit */
      select (b_cred) ;
            when ('CR')          b_cred1=1;
            when ('KI')          b_cred2=1;
      otherwise;
     end;

mjneg1=0; mjneg2=0;       /*Major negative bureau */
      select (mjneg) ;
            when ('CR')          mjneg1=1;
            when ('KI')          mjneg2=1;
      otherwise;
     end;

benq0=0; benq1=0;         /*Number of bureau enquiries*/
      select (benq) ;
            when (0)            benq0=1;
            when (99)           benq1=1;
      otherwise;
     end;
```

```
b_1b0=0; b_1b1=0;              /*Bureau quantity 1*/
    select (b_1b) ;
        when (0)               b_1b0=1;
        when (99)              b_1b1=1;
    otherwise;
   end;

b_2b0=0; b_2b1=0;              /*Bureau quantity 2*/
    select (b_2b) ;
        when (0)               b_2b0=1;
        when (99)              b_2b1=1;
    otherwise;
   end;

b_3b0=0; b_3b1=0; b_3b2=0;          /*Bureau quantity 3*/
    select;
        when (b_3b=0)          b_3b0=1;
        when (b_3b=1)          b_3b1=1;
        when (b_3b=99)         b_3b2=1;
    otherwise;
   end;

b_4b0=0; b_4b1=0; b_4b2=0; b_4b3=0;       /*Bureau quantity 4*/

    select;
        when (b_4b=0)          b_4b0=1;
        when (b_4b=1)          b_4b1=1;
        when (1<b_4b<9)        b_4b2=1;
        when (b_4b=99)         b_4b3=1;
    otherwise;
   end;

totc0=0; totc1=0;              /*Number of credit repaid*/

    select (totc);
        when (0)               totc0=1;
        when (99)              totc1=1;
    otherwise;
   end;

bliv0=0; bliv1=0; bliv2=0;           /*Number of credit open */

    select (bliv);
        when (0)               bliv0=1;
        when (1,2,3,4)         bliv1=1;
        when (99)              bliv2=1;
    otherwise;
   end;
```

```
mnnk0=0; mnnk1=0;            /*Minor negative bureau*/
 select (mnnk);
         when (0)      mnnk0=1;
         when (99)      mnnk1=1;
     otherwise;
     end;
```

A20. Binary variable coding for the generic model

```
phone1=0; phone2=0; phone3=0;      /*Home phone given*/

select (phone);
     when ('1','Y')              phone1=1;
     when ('0','N')              phone2=1;
     when ('3','M')              phone3=1;
     otherwise;  /*secret number */
end;

    rstat1=0; rstat2=0; rstat3=0; rstat4=0; rstat5=0;  /*Residential status*/
select (rstat);

     when ('1','6')      rstat1=1;
     when ('2')          rstat2=1;
     when ('3')          rstat3=1;
     when ('4')          rstat4=1;
     when ('5')          rstat5=1;
     otherwise;  /* other + houseboat*/
end;

     mstat1=0; mstat2=0; mstat3=0;               /*Marital status*/
select (mstat);
     when ('1')          mstat1=1;
     when ('2')          mstat2=1;
     when ('3','4')      mstat3=1;
     otherwise;  /* living together */
end;

cardi1=0;
select (cardi);
when ('1')            cardi1=1;
     otherwise;   /* insurance */
end;
```

```
    credi0=0;                    /*Credit insurance */
select (credi);
    when ('0')    credi0=1;
otherwise;   /* insurance */
end;

    pdate0=0; pdate1=0; pdate2=0; pdate3=0;                  /*Payment date */
select (pdate);
    when ('00')         pdate0=1;
    when ('01')         pdate1=1;
    when ('08')         pdate2=1;
    when ('15')         pdate3=1;
    otherwise;    /* 22, 28, etc */
end;

    bus1=0; bus2=0; bus3=0; bus4=0; bus5=0; bus6=0;    /*Type of business */
select (bus);
    when ('63','62','55','51')                          bus1=1;
    when ('09','12','03','17')                          bus2=1;
    when ('11','14','27','52','13','21','18','10')      bus3=1;
    when ('23','26','01')                               bus4=1;
    when ('24','19','56','25','57','54','16','05','15') bus5=1;
    when ('04','20','02','08','06','22','66','64','07') bus6=1;
    otherwise;   /* 99, ' ', etc*/
end;

    gcode1=0; gcode2=0; gcode3=0; gcode4=0; gcode5=0; gcode6=0; gcode7=0;
select (gcode);                         /*Goods code*/
    when ('12')                              gcode1=1;
    when ('19','43','20','56','85')          gcode2=1;
    when ('00','05','58')                    gcode3=1;
    when ('59','48','25','06','01','04','03') gcode4=1;
    when ('17','77','99','10','08','26')     gcode5=1;
    when ('66','09','14','11','07')          gcode6=1;
    when ('95','13','97','16','31')          gcode7=1;
    otherwise;  /* total < 100 */
end;

    spage0=0; spage1=0; spage2=0; spage3=0; spage4=0; spage5=0;
select;                       /*Spouse's age*/
    when (spage=999)          spage1=1;
    when (spage<17)           spage0=1;
    when (16<spage<22)        spage2=1;
    when (21<spage<28)        spage3=1;
    when (27<spage<41)        spage4=1;
    when (40<spage<61)        spage5=1;
    otherwise;   /* 61 + */
end;
```

```
    gp0=0; gp1=0; gp2=0; gp3=0; gp4=0; gp5=0;  /*Goods price */
select;
     when (gp=0)              gp0=0;
    when (gp<200)            gp1=1;
    when (199<gp<400)        gp2=1;
    when (399<gp<600)        gp3=1;
     when (599<gp<800)       gp4=1;
    when (799<gp<1000)       gp5=1;
    otherwise;    /* 1000+ */
end;

    tad1=0; tad2=0; tad3=0; tad4=0; tad5=0; tad6=0;        /*Time at address */
select;
    when (0<tad<5)             tad1=1;
    when (4<tad<107)           tad2=1;
    when (106<tad<407)         tad3=1;
    when (406<tad<707)         tad4=1;
    when (706<tad<1500)        tad5=1;
    when (1411<tad<9999)       tad6=1;
    otherwise;   /* tad=0 */
end;

    toj0=0; toj1=0; toj2=0; toj3=0; toj4=0; toj5=0;           /*Time in employment */
select;
     when (0<=toj<7)                                            toj1=1;
     when (6<toj<300)                                           toj2=1;
     when (211<toj<700)                                         toj3=1;
     when (611<toj<1300)                                        toj4=1;
     when (toj=9992  or toj=9995 or toj=9998 or toj=9999)       toj5=1;
     when (toj=.)                                               toj0=1;
     otherwise;        /*more than 13 years */
 end;

    occup1=0; occup2=0; occup3=0; occup4=0; occup5=0;          /*Occupation */
select (occup);
    when (' ','42','41','06','40')               occup1=1;
    when ('02')                                  occup2=1;
    when ('07')                                  occup3=1;
    when ('01')                                  occup4=1;
    when ('08','20','31','33','30','05','32')    occup5=1;
    otherwise;    /* widow rent , housewife, retired */
end;

ephone1=0;                     /*Employer's phone given */
select (ephone);
    when ('1')    ephone1=1;
    otherwise;  /*not given */ end;
```

```
age1=0; age2=0; age3=0; age4=0;              /*Applicant's age */
select;
     when (23<age<30)        age1=1;
     when (29<age<36)        age2=1;
     when (35<age<51)        age3=1;
     when (50<age)           age4=1;
     otherwise;     /* <24 */
end;

    kids1=0; kids2=0; kids3=0;               /*Number of dependants */
select;
     when (kids=1)           kids1=1;
     when (kids=2)           kids2=1;
     when (2<kids<99)        kids3=1;
     otherwise;    /* 0,99 */
end;

stateh=0; stateg=0;                          /*Country*/
select (country);
     when ('G')         stateg=1;
     when ('H')         stateh=1;
 otherwise;    /* Belgium */
end
```

A21. Parameter estimates of logistic regression models

(Main effects/Binary/Stepwise for Belgium, Germany, generic; Interactions (1)/Stepwise for the Netherlands)

	Belgium		The Netherlands		Germany		Generic –no country		Generic – country	
Characteristics	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate
Age	AGE1	-0.28	AGE1	-0.91	AGE1	-0.98	AGE1	0.36	AGE1	0.38
	AGE5	0.21	AGE2	-0.68	AGE2	-0.81	AGE2	0.54	AGE2	0.55
			AGE3	-0.44	AGE3	-0.52	AGE3	0.66	AGE3	0.66
			AGE4	-0.36	AGE4	-0.32	AGE4	0.95	AGE4	0.97
			AGE5	-0.23						
Type of business	BUS11	-0.19	BTYPE2	-0.15	BTYPE1	-0.32	BUS2	-0.20	BUS2	-0.20
	BUS14	0.23	BTYPE4	0.16	BTYPE2	-0.18	BUS5	0.11	BUS5	0.12
	BUS15	0.24	BTYPE5	0.26	BTYPE4	0.09	BUS6	0.34	BUS6	0.36
	BUS16	0.65	BTYPE6	1.09	BTYPE5	0.30				
					BTYPE6	0.54				
Credit/card insurance	CREDI0	0.21	CREDI0	0.31			CREDI0	0.29	CREDI0	0.33
					CARDI1	-0.39	CARDI1	-0.14		
Employer's phone	EPHONE1	0.12								

	Belgium		The Netherlands		Germany		Generic –no country		Generic – country	
Characteristics	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate
Goods code	GDSC1	-0.60	GCPD11	-1.32	GCODE1	-0.21	GCODE1	-1.37	GCODE1	-1.53
	GDSC2	-0.62	GCPD14	-2.05	GCODE4	0.14	GCODE2	-0.72	GCODE2	-0.72
	GDSC3	-0.32	GCPD21	-0.46	GCODE5	0.35	GCODE3	-0.37	GCODE3	-0.41
	GDSC6	0.36	GCPD24	-1.01			GCODE4	-0.23	GCODE4	-0.25
			GCPD41	0.25					GCODE5	-0.11
Payment date	PDATE14	-0.67	GCPD42	0.48	PDAY1	0.24	PDATE0	-0.45	PDATE0	-0.53
	PDATE15	0.13	GCPD51	0.40			PDATE1	0.36	PDATE1	0.29
	PDATE8	0.38	GCPD52	0.56			PDATE2	0.37	PDATE2	0.41
			GCPD61	0.40			PDATE3	0.18	PDATE3	0.18
			GCPD71	0.11						
Goods price	GP1	0.56	GPR1	-0.11					GP1	0.21
	GP2	0.46	GPR3	-0.22			GP2	0.09	GP2	0.18
	GP3	0.53	GPR4	-0.34					GP3	0.11
	GP4	0.21					GP5	-0.23	GP5	-0.17
	GP5	0.33								
	GP6	-0.24								
Number of dependants	KIDS13	-0.55	KIDS1	-0.59	KIDS3	-0.25			KIDS1	0.09
							KIDS3	-0.37	KIDS3	-0.30

	Belgium		The Netherlands		Germany		Generic –no country		Generic – country	
Characteristics	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate
Marital status			MSTAT1	0.34	MSTAT1	0.14	MSTAT1	0.23	MSTAT1	0.11
			MSTAT2	0.32	MSTAT2	-0.15			MSTAT2	-0.13
			MSTAT3	0.59			MSTAT3	0.14		
Occupation	OCCUP3	-0.44	OCCUP1	-0.82	OCC1	0.52	OCCUP2	-0.60	OCCUP2	-0.60
	OCCUP4	0.73	OCCUP2	-0.65	OCC3	-0.60	OCCUP5	0.20	OCCUP5	0.18
			OCCUP3	-0.12						
Home phone	PHONE1	0.39	PHONE1	0.80	PHONE1	-0.81	PHONE1	0.70	PHONE1	0.72
Residential status	RSTAT1	0.47	RSTAT3	-0.23	RSTAT1	0.25	RSTAT1	0.23	RSTAT1	0.26
	RSTAT3	0.55	RSTAT4	0.41	RSTAT3	-0.13	RSTAT2	-0.29	RSTAT2	-0.28
							RSTAT3	-0.21	RSTAT3	-0.22
							RSTAT4	-0.29	RSTAT4	-0.33
Spouse's age	SPAGE1	-0.32	SPAGE1	-0.24	SPAGE0	0.10	SPAGE1	-0.29	SPAGE1	-0.27
	SPAGE2	-0.35	SPAGE4	0.21	SPAGE1	-0.74	SPAGE2	-0.31	SPAGE2	-0.31
					SPAGE4	0.25				
					SPAGE5	0.52				
Time at address	TAD1	-0.73	TAD1	0.17	TAD1	-0.54	TAD1	-0.35	TAD1	-0.37
	TAD2	-0.43	TAD2	0.27	TAD2	-0.29	TAD2	-0.10	TAD2	-0.12
	TAD3	-0.61	TAD3	0.35	TAD3	-0.14	TAD6	0.27	TAD6	0.26
	TAD4	-0.28								

	Belgium		The Netherlands		Germany		Generic –no country		Generic – country	
Characteristics	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate
Time in employment	TOJ1	-0.38	TOJ1	-0.34	TOJ1	-0.76	TOJ1	-0.57	TOJ1	-0.58
	TOJ2	-0.17	TOJ2	-0.16	TOJ2	-0.51	TOJ2	-0.30	TOJ2	-0.30
	TOJ5	0.37			TOJ3	-0.31	TOJ3	-0.14	TOJ3	-0.14
	TOJ6	0.49			TOJ6	-0.57	TOJ5	-0.46	TOJ5	-0.42
Country									STATEH	-0.33
									STATEG	-0.12

A22. Collinearity diagnostics of logistic regression models. Variance inflation factor.

BELGIUM			NETHERLANDS			GERMANY			GENERIC		
Variable	DF	Variance Inflation	Variable	DF	Variance Inflation	Variable	DF	Variance Inflation	Variable	DF	Variance Inflation
PHONE1	1	1.01470447	PHONE1	1	1.06157133	RSTAT1	1	1.95272085	PHONE1	1	1.04332584
RSTAT1	1	1.27926495	RSTAT3	1	1.31744779	RSTAT3	1	1.93074226	RSTAT1	1	2.03447437
RSTAT3	1	2.85304064	RSTAT4	1	1.33272977	MSTAT1	1	3.12366833	RSTAT2	1	1.97720036
OCCUP3	1	1.06467426	MSTAT1	1	1.37412141	MSTAT2	1	3.32657068	RSTAT3	1	2.26507692
OCCUP4	1	1.19380093	MSTAT2	1	1.11393709	PHONE1	1	1.04646744	RSTAT4	1	1.05517150
AGE1	1	1.27275841	MSTAT3	1	1.66041656	KIDS3	1	1.04934642	MSTAT1	1	3.26008261

BELGIUM			NETHERLANDS			GERMANY			GENERIC		
Variable	DF	Variance Inflation	Variable	DF	Variance Inflation	Variable	DF	Variance Inflation	Variable	DF	Variance Inflation
AGE5	1	1.08632840	BTYPE2	1	1.06876429	BTYPE1	1	1.03997790	MSTAT3	1	1.40829870
BUS11	1	1.17072362	BTYPE4	1	1.17725723	BTYPE2	1	1.13240845	CARDI1	1	1.07430772
BUS14	1	1.15117590	BTYPE5	1	1.08938693	BTYPE4	1	1.19750384	CREDI0	1	1.30701620
BUS15	1	1.16671732	BTYPE6	1	1.14676908	BTYPE5	1	1.11144542	PDATE0	1	3.38043831
BUS16	1	1.05184741	SPAGE1	1	1.18844499	BTYPE6	1	1.04138719	PDATE1	1	2.56514545
KIDS13	1	1.03065618	SPAGE4	1	1.41598972	OCC1	1	1.04235839	PDATE2	1	1.85903836
GDSC1	1	1.14908251	CREDI0	1	1.37375648	OCC3	1	1.03726990	PDATE3	1	2.12972165
GDSC2	1	1.39346470	TAD1	1	1.88478513	CARDI1	1	1.00510972	BUS2	1	1.07932505
GDSC3	1	1.12039777	TAD2	1	2.06798177	AGE1	1	2.01630761	BUS5	1	1.08750095
GDSC6	1	1.20921799	TAD3	1	1.96723556	AGE2	1	1.70053606	BUS6	1	1.08167782
GP1	1	1.65100431	KIDS1	1	1.00666086	AGE3	1	1.61273136	GCODE1	1	2.68017979
GP2	1	1.43921411	GPR1	1	1.51708886	AGE4	1	1.23009215	GCODE2	1	1.30134080
GP3	1	1.16583053	GPR3	1	1.80755367	SPAGE0	1	1.61096513	GCODE3	1	1.34810394
GP6	1	1.31977254	GPR4	1	1.88968146	SPAGE1	1	1.01836522	GCODE4	1	1.27208407
SPAGE1	1	1.35958160	OCCUP1	1	1.41688023	SPAGE4	1	1.43759172	SPAGE1	1	2.76021358
SPAGE2	1	1.17787959	OCCUP2	1	1.27669224	SPAGE5	1	1.10107523	SPAGE2	1	1.12984038
CREDI0	1	1.04090146	OCCUP3	1	1.77200111	TAD1	1	1.65502687	GP2	1	1.14024375
PDATE8	1	1.32327431	TOJ1	1	1.25491458	TAD2	1	1.63543107	GP5	1	1.05703330

BELGIUM			NETHERLANDS			GERMANY			GENERIC		
Variable	DF	Variance Inflation	Variable	DF	Variance Inflation	Variable	DF	Variance Inflation	Variable	DF	Variance Inflation
PDATE14	1	1.13958827	TOJ2	1	1.28308812	TAD3	1	1.29575621	TAD1	1	1.11139261
PDATE15	1	1.31685231	AGE1	1	2.08596145	TOJ1	1	1.53718972	TAD2	1	1.12573458
TAD1	1	3.54143666	AGE2	1	2.57850149	TOJ2	1	1.65655748	TAD6	1	1.22102961
TAD2	1	1.88428702	AGE3	1	2.34745255	TOJ3	1	1.29502405	TOJ1	1	1.63522676
TAD3	1	1.32223932	AGE4	1	2.96985367	TOJ6	1	1.41728102	TOJ2	1	1.95815573
TAD4	1	1.56159663	AGE5	1	2.84850411	GCODE1	1	1.17320396	TOJ3	1	1.62500433
TOJ1	1	1.15627142	GCPD11	1	1.07957378	GCODE4	1	1.48678620	TOJ5	1	1.37379687
TOJ2	1	1.14289935	GCPD14	1	2.17509126	GCODE5	1	1.53677452	OCCUP2	1	1.05442607
TOJ5	1	1.07107749	GCPD21	1	1.10162638	PDAY1	1	1.03264399	OCCUP5	1	1.30765032
TOJ6	1	1.26599692	GCPD24	1	1.63667461				EPHONE1	1	1.74307015
EPHONE1	1	1.15098363	GCPD41	1	1.31989851				AGE1	1	1.83516636
			GCPD42	1	1.02809074				AGE2	1	1.98607396
			GCPD51	1	1.60270840				AGE3	1	2.76384575
			GCPD52	1	1.07212316				AGE4	1	2.05839048
			GCPD61	1	1.24856611				KIDS3	1	1.11106571
			GCPD71	1	1.42179950						

A23. Parameter estimates of logistic regression models developed on all information available.

Belgium		The Netherlands		Germany	
Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate
INTERCEP	0.83	INTERCEP	0.27	INTERCEP	1.79
EPHONE1	0.13				
PHONE2	-0.52	PHONE1	0.78	PHONE1	-0.74
RSTAT1	0.47	RSTAT1	0.24	RSTAT2	-0.17
RSTAT3	0.51	RSTAT2	0.16	RSTAT3	-0.21
		RSTAT4	0.56		
MSTAT1	0.16	MSTAT1	0.33	MSTAT1	0.23
		MSTAT2	0.38	MSTAT2	-0.12
		MSTAT3	0.60		
OCCUP3	-0.49	OCCUP1	-0.74	OCC3	-0.51
OCCUP4	0.70	OCCUP2	0.20		
BUS11	-0.17	BTYPE1	-0.53	BTYPE1	-0.32
BUS14	0.23	BTYPE2	-0.14	BTYPE2	-0.34
BUS15	0.23	BTYPE4	0.14	BTYPE3	-0.13
BUS16	0.58	BTYPE5	0.24	BTYPE5	0.13
		BTYPE6	1.14	BTYPE6	0.43
		AGE1	0.32	AGE1	-0.47
		AGE2	0.60	AGE2	-0.55
		AGE3	0.67	AGE3	-0.39
		AGE4	0.68	AGE4	-0.29
		AGE5	0.78		
		AGE6	0.97		
KIDS13	-0.53			KIDS2	0.10
SPAGE2	-0.25	SPAGE1	-0.22	SPAGE1	-0.70
		SPAGE4	0.23		
		CREDI0	0.27	CARDI1	-0.38
CREDI0	0.23	CREDI1	-0.23		
GP6	-0.37	GPR3	-0.17	GPRICE3	-0.10
		GPR4	-0.35		

Belgium		The Netherlands		Germany	
Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate
GDSC1	-0.45	GCPD21	-0.34	GCODE1	-0.15
GDSC2	-0.55	GCPD24	1.13	GCODE4	0.14
GDSC3	-0.28	GCPD31	-0.17	GCODE5	0.25
GDSC6	0.27	GCPD33	-0.31		
PDATE8	0.40	GCPD51	0.18	PDAY1	0.41
PDATE15	0.15	GCPD52	0.48	PDAY2	0.16
TAD1	-0.37	TAD1	0.16	TAD1	-0.29
TAD3	-0.36	TAD2	0.24	TAD2	-0.09
		TAD3	0.29		
TOJ1	-0.23	TOJ1	-0.31	TOJ1	-0.50
TOJ5	0.27	TOJ2	-0.12	TOJ2	-0.31
TOJ6	0.34			TOJ3	-0.19
				TOJ6	-0.48
Additional variables					
TADDR1	0.14	LOAN3	0.15	NAT1	0.30
CARD1	0.32	LOAN4	0.12	ININST4	0.09
CARD2	0.43	INSTAL0	-0.29	L_CR2	-0.17
PREMP0	0.18	NEGBU3	-0.87	L_CR3	-0.35
TBANK1	-1.18	LIVFT0	0.11	TBANK1	-0.97
TBANK2	-1.02	LIVECR0	0.10	TBANK2	-0.81
TBANK3	-0.76	LASTA0	-0.38	TBANK3	-0.56
TBANK4	-0.64	LASTA4	0.17	TBANK4	-0.33
TBANK5	-0.43	LASTA5	0.20	TBANK5	-0.12
TBANK6	-0.34	F_BURO2	-0.28	BTIM5	-0.20
ININST1	0.28	F_BURO4	-0.26	T1B3	-0.28
RETAIL1	0.30	F_BURO5	-0.19	T1B4	0.16
RETAIL2	0.23	LASTAP2	-1.63	T1B5	0.39
AGREEM1	0.32	LASTAP3	-1.36	T2B1	-0.43
AGREEM2	0.32	DEAL2	0.11	BSCOD3	0.27
SPOCC1	0.28	BANKC1	0.31	BSCOD4	0.17
SPOCC2	0.29	BANKC2	0.42	S_DL1	-0.44
		BANKC3	0.28	S_DL2	-0.30
		BANKC4	0.23	S_DL3	-0.22
		NATION1	0.33	THINST1	0.24

Belgium		The Netherlands		Germany	
Variable	Parameter Estimate	Variable	Parameter Estimate	Variable	Parameter Estimate
		AGR1	-2.67	PDEP0	-0.43
		AGR2	-2.80	PDEP1	-0.56
		AGR3	-0.78	L_INIT2	-0.17
		AGR4	-0.45	UNDER1	0.13
		AGR5	-0.44	SPOUSE0	-0.11
				CRCARD1	0.74
				CRCARD2	0.41
				E_W1	0.21
				TYPEB1	0.46
				B_CRED1	-0.16
				B_1B0	0.64
				B_3B0	-0.27
				B_4B1	0.08

A24. Parameter estimates of PH model with variable-by-time interactions.

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
PHONE1	1	-0.289156	0.38408	0.56677	0.4515	0.749
PHONE2	1	0.616004	0.39243	2.46395	0.1165	1.852
PHONE3	1	0.335587	0.39433	0.72424	0.3948	1.399
RSTAT1	1	-0.096635	0.12169	0.63062	0.4271	0.908
RSTAT2	1	0.299970	0.11466	6.84469	0.0089	1.350
RSTAT3	1	0.297436	0.11652	6.51613	0.0107	1.346
RSTAT4	1	0.167680	0.22715	0.54494	0.4604	1.183
RSTAT5	1	0.022395	0.12812	0.03055	0.8612	1.023
MSTAT2	1	0.243539	0.11792	4.26530	0.0389	1.276
MSTAT3	1	0.133539	0.09930	1.80857	0.1787	1.143
CARDI0	1	0.319703	0.12431	6.61379	0.0101	1.377
CARDI1	1	0.647205	0.14583	19.69776	0.0001	1.910
CREDI0	1	-0.500802	0.36247	1.90896	0.1671	0.606
CREDI1	1	-0.001944	0.37818	0.0000264	0.9959	0.998
CREDI2	1	0.069712	0.49254	0.02003	0.8874	1.072
CREDI3	1	-0.224565	0.37277	0.36292	0.5469	0.799
PDATE0	1	0.477141	0.15410	9.58773	0.0020	1.611
PDATE1	1	-0.573815	0.08353	47.18759	0.0001	0.563
PDATE2	1	-0.276310	0.09275	8.87520	0.0029	0.759
PDATE3	1	-0.197395	0.08331	5.61349	0.0178	0.821
BUS1	1	-0.216142	0.26851	0.64799	0.4208	0.806
BUS2	1	0.213487	0.13547	2.48337	0.1151	1.238
BUS3	1	0.161330	0.13173	1.49984	0.2207	1.175
BUS4	1	0.112943	0.12463	0.82127	0.3648	1.120
BUS5	1	-0.104166	0.13921	0.55988	0.4543	0.901
BUS6	1	-0.226631	0.13274	2.91502	0.0878	0.797
GCODE1	1	1.354146	0.30526	19.67910	0.0001	3.873
GCODE2	1	0.861345	0.28571	9.08874	0.0026	2.366
GCODE3	1	0.371956	0.28107	1.75131	0.1857	1.451
GCODE4	1	0.230121	0.27852	0.68265	0.4087	1.259
GCODE5	1	0.046510	0.28539	0.02656	0.8705	1.048
GCODE6	1	-0.198110	0.27969	0.50171	0.4787	0.820
GCODE7	1	-0.142556	0.29853	0.22804	0.6330	0.867
SPAGE0	1	1.139280	0.41805	7.42678	0.0064	3.125
SPAGE1	1	1.494603	0.41194	13.16357	0.0003	4.458
SPAGE2	1	1.451881	0.43514	11.13289	0.0008	4.271
SPAGE3	1	1.051538	0.41867	6.30834	0.0120	2.862
SPAGE4	1	1.008686	0.41000	6.05273	0.0139	2.742
SPAGE5	1	1.205015	0.40805	8.72103	0.0031	3.337
GP1	1	0.045755	0.11795	0.15048	0.6981	1.047
GP2	1	-0.214309	0.07907	7.34555	0.0067	0.807
GP3	1	-0.213211	0.07831	7.41318	0.0065	0.808

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
GP4	1	0.005139	0.08489	0.00366	0.9517	1.005
GP5	1	0.246872	0.08742	7.97436	0.0047	1.280
TAD1	1	0.346535	0.11938	8.42561	0.0037	1.414
TAD2	1	0.126030	0.10250	1.51194	0.2188	1.134
TAD3	1	-0.046283	0.09746	0.22554	0.6349	0.955
TAD4	1	-0.161948	0.11420	2.01111	0.1562	0.850
TAD5	1	-0.247231	0.11730	4.44251	0.0351	0.781
TAD6	1	-0.485837	0.14409	11.36895	0.0007	0.615
TOJ0	1	-0.060962	0.27235	0.05010	0.8229	0.941
TOJ1	1	0.484413	0.11118	18.98454	0.0001	1.623
TOJ2	1	0.209136	0.10226	4.18294	0.0408	1.233
TOJ3	1	0.133371	0.10542	1.60071	0.2058	1.143
TOJ4	1	-0.039505	0.11336	0.12145	0.7275	0.961
TOJ5	1	0.439647	0.19420	5.12498	0.0236	1.552
OCCUP1	1	-0.149708	0.29252	0.26192	0.6088	0.861
OCCUP2	1	0.310002	0.21362	2.10601	0.1467	1.363
OCCUP3	1	-0.034729	0.26173	0.01761	0.8944	0.966
OCCUP4	1	-0.269206	0.19227	1.96046	0.1615	0.764
OCCUP5	1	-0.445539	0.19740	5.09398	0.0240	0.640
EPHONE1	1	0.077853	0.07222	1.16215	0.2810	1.081
A1	1	-0.192912	0.07499	6.61740	0.0101	0.825
A2	1	-0.218345	0.09248	5.57373	0.0182	0.804
A3	1	-0.458729	0.09835	21.75602	0.0001	0.632
A4	1	-0.590038	0.14838	15.81203	0.0001	0.554
K1	1	-0.152477	0.09209	2.74125	0.0978	0.859
K2	1	-0.271019	0.11004	6.06647	0.0138	0.763
K3	1	0.355878	0.12820	7.70580	0.0055	1.427
PHONT1	1	0.020226	0.03563	0.32227	0.5702	1.020
PHONT2	1	-0.007988	0.03656	0.04774	0.8270	0.992
PHONT3	1	0.018073	0.03676	0.24177	0.6229	1.018
RSTATT1	1	-0.007888	0.01106	0.50838	0.4758	0.992
RSTATT2	1	-0.001571	0.01047	0.02253	0.8807	0.998
RSTATT3	1	-0.009050	0.01069	0.71635	0.3973	0.991
RSTATT4	1	0.009054	0.02116	0.18315	0.6687	1.009
RSTATT5	1	-0.001620	0.01184	0.01870	0.8912	0.998
MSTATT2	1	-0.004805	0.01073	0.20067	0.6542	0.995
MSTATT3	1	-0.009135	0.00888	1.05819	0.3036	0.991
CARDIT0	1	-0.024107	0.01112	4.70341	0.0301	0.976
CARDIT1	1	-0.046728	0.01335	12.25006	0.0005	0.954
CREDIT0	1	-0.002101	0.03548	0.00351	0.9528	0.998
CREDIT1	1	-0.024462	0.03685	0.44072	0.5068	0.976
CREDIT2	1	-0.031334	0.05568	0.31669	0.5736	0.969
CREDIT3	1	-0.006766	0.03629	0.03477	0.8521	0.993
PDATET0	1	0.002728	0.01469	0.03450	0.8527	1.003
PDATET1	1	0.026674	0.00764	2.18529	0.0005	1.027
PDATET2	1	-0.004846	0.00877	0.30556	0.5804	0.995

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
PDATET3	1	0.004916	0.00776	0.40154	0.5263	1.005
BUST1	1	0.014249	0.02352	0.36712	0.5446	1.014
BUST2	1	0.006313	0.01229	0.26402	0.6074	1.006
BUST3	1	-0.001703	0.01197	0.02024	0.8869	0.998
BUST4	1	-0.001301	0.01130	0.01326	0.9083	0.999
BUST5	1	0.009887	0.01250	0.62517	0.4291	1.010
BUST6	1	0.001336	0.01200	0.01240	0.9113	1.001
GCODET1	1	-0.050798	0.02654	3.66292	0.0556	0.950
GCODET2	1	-0.050276	0.02433	4.27165	0.0388	0.951
GCODET3	1	-0.030370	0.02371	1.64090	0.2002	0.970
GCODET4	1	-0.029486	0.02344	1.58213	0.2085	0.971
GCODET5	1	-0.022831	0.02410	0.89775	0.3434	0.977
GCODET6	1	-0.012440	0.02350	0.28023	0.5966	0.988
GCODET7	1	-0.014943	0.02525	0.35012	0.5540	0.985
SPAGET0	1	-0.030341	0.03437	0.77909	0.3774	0.970
SPAGET1	1	-0.040058	0.03377	1.40712	0.2355	0.961
SPAGET2	1	-0.025580	0.03614	0.50111	0.4790	0.975
SPAGET3	1	-0.012838	0.03437	0.13950	0.7088	0.987
SPAGET4	1	-0.014184	0.03352	0.17902	0.6722	0.986
SPAGET5	1	-0.040646	0.03336	1.48453	0.2231	0.960
GPT1	1	-0.020082	0.01179	2.90076	0.0885	0.980
GPT2	1	0.008307	0.00713	1.35813	0.2439	1.008
GPT3	1	0.013496	0.00693	3.79689	0.0513	1.014
GPT4	1	-0.000353	0.00772	0.00209	0.9635	1.000
GPT5	1	-0.007866	0.00796	0.97667	0.3230	0.992
TADT1	1	-0.015853	0.01116	2.01850	0.1554	0.984
TADT2	1	-0.014720	0.00948	2.41180	0.1204	0.985
TADT3	1	-0.006648	0.00896	0.55030	0.4582	0.993
TADT4	1	0.002782	0.01039	0.07172	0.7889	1.003
TADT5	1	0.005565	0.01059	0.27593	0.5994	1.006
TADT6	1	0.008331	0.01284	0.42097	0.5165	1.008
TOJT0	1	-0.022853	0.02571	0.79017	0.3740	0.977
TOJT1	1	-0.003059	0.01001	0.09332	0.7600	0.997
TOJT2	1	0.002830	0.00909	0.09687	0.7556	1.003
TOJT3	1	-0.004337	0.00939	0.21335	0.6442	0.996
TOJT4	1	-0.003617	0.01005	0.12955	0.7189	0.996
TOJT5	1	-0.011756	0.01764	0.44414	0.5051	0.988
OCCUPT1	1	0.018025	0.02656	0.46064	0.4973	1.018
OCCUPT2	1	-0.004110	0.01919	0.04587	0.8304	0.996
OCCUPT3	1	-0.003997	0.02431	0.02704	0.8694	0.996
OCCUPT4	1	0.003287	0.01718	0.03660	0.8483	1.003
OCCUPT5	1	0.010420	0.01759	0.35073	0.5537	1.010
EPHONET1	1	-0.008690	0.00659	1.73797	0.1874	0.991

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
AT1	1	-0.007266	0.00701	1.07593	0.2996	0.993
AT2	1	-0.020220	0.00866	5.45781	0.0195	0.980
AT3	1	-0.005508	0.00904	0.37103	0.5424	0.995
AT4	1	-0.011448	0.01365	0.70315	0.4017	0.989
KT1	1	0.007479	0.00838	0.79633	0.3722	1.008
KT2	1	0.025470	0.00952	7.15363	0.0075	1.026
KT3	1	-0.007841	0.01176	0.44436	0.5050	0.992

A25. Hazard ratios for models with different levels of information. Belgium

Variable	Model								
	Personal	Personal, Purchase	Personal, Purchase, ATS_1	Personal, Purchase, ATS_{t-1}	Personal, Purchase, ATS_{t-1}, ATS_{t-2}	From 2nd period- more info	Personal, Purchase, ATS_{t-1} – Goods	Personal, Purchase, ATS_{t-1} -Bads after 2nd p	Personal, Purchase, ATS_{t-1} – Bads before 2 p
	1	2	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Phone number given								1.302	1.850
Kids: 1 child									1.390
Marital: Widowed	1.118	1.125	1.144						
Res: Renting room, parents	1.090	1.082	1.090		1.110	1.127	1.120		
Res; Renting house/flat	1.109	1.088	1.101	1.085	1.147	1.174	1.152		
Occup: Retired		0.925							
Occup: Part-time	1.111								
Occup: Self-employed	0.798	0.885	0.903	0.881	0.876			0.628	
Type of business – Unknown	1.149	1.093			1.129				
Type of business – 21	1.206	1.122	1.086	1.115	1.158	1.124	1.080		
Age : under 21		0.879							
Age : 22-27						1.071			
No spouse		0.926	0.934		0.942			0.821	
Time at address : 6 months	1.063		1.063		1.126	1.146	1.090	0.873	
Time at address : 6 mths – 1 yr	1.110	1.054	1.111	1.046	1.163	1.181	1.123		
Time on job: 1 yr						1.064			
Allowance			1.128	1.092		1.191	1.146		
No card insurance		0.926	0.925	0.941	0.899	0.883	0.921		
No credit insurance		1.129	1.092		1.136	1.136	1.111		

	1	2	3	4	5	6	7	8	9
Product type- computers		0.888		0.890	0.805	0.818		1.324	
Product type- TV					0.941				
Product type- household1									1.440
Product type- household2									1.446
Product type- phones		1.058	1.048	1.057		1.076	1.094		
Price = 0		1.324		1.161	1.802	1.621			
Price < =10,000 BEF		1.878	1.231	1.628	2.026	2.003	1.298		1.407
10,000 < Price <= 16,000 BEF		1.627	1.089	1.444	1.697	1.602	1.116		
16,000 < Price <= 20,000 BEF		1.388		1.251	1.372	1.315			
20,000 < Price <= 40,000 BEF		1.227		1.166	1.279	1.255	1.059		
Pay date 01		0.727	0.731	0.741			0.721	0.687	
Pay date 08		1.184	1.146	1.171	1.397	1.389	1.142		0.666
Pay date 14,15		1.450	1.421	1.433	1.247	1.229	1.438	1.328	
Agreement type – budget		2.068	2.349	2.171	1.571	1.963	2.406	3.350	
ATS $_{t-1}$ – over credit limit			0.334	0.021	0.010	0.079	0.331	0.286	
ATS $_{t-1}$ = 0			0.790	1.775	1.867	3.399	0.842	0.532	
ATS $_{t-1}$ <= 5000 BEF			0.655	0.727	0.315	0.637	0.699	0.575	
5000< ATS $_{t-1}$ < =10,000 BEF			0.801		0.730		0.823	0.737	
10,000< ATS $_{t-1}$ <= 20,000 BEF			0.880				0.870		1.326
ATS $_{t-2}$ – over credit limit					2.406				
ATS $_{t-2}$ = 0					2.601				
ATS $_{t-2}$ <= 5000 BEF					1.677				
5000< ATS $_{t-2}$ < =10,000 BEF					1.153				
1 missed payment						0.008			
2 missed payment						0.013			

The Netherlands

Variable	Model								
	Per-sonal	Personal, Purchase	Personal, Purchase, ATS_1	Personal, Purchase, ATS_{t-1}	Personal, Purchase, ATS_{t-1}, ATS_{t-2}	From 2nd period- more info	Personal, Purchase, ATS_{t-1} – Goods	Personal, Purchase, ATS_{t-1} -Bads after 2nd p	Personal, Purchase, ATS_{t-1} – Bads before 2 p
	1	2	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Phone given		1.104	1.094	1.068	1.201				1.356
Kids – no info	1.288								
Marital: Married					1.088				
Marital: Single, Widowed		0.943	0.942	0.954	0.903	0.887			
Res: Rented house, flat	1.196								
Res: Rented room	1.108	1.068	1.070						
Res: Parents	1.096	1.118	1.097	1.095	1.102	1.073	1.064	1.125	
Occup: Self-emp	0.812	0.806	0.853	0.850	0.854				
Occup: Full-time		0.936							
Type of business: Catering, shop	1.093	1.067	1.067	1.059			1.054		
Type of business: Benefit, agency	1.132								
Age: under 22		0.795	0.800	0.834	0.810				
Age: 22 - 31	0.938	0.867	0.863	0.875	0.930				
Age: 32-36		0.900	0.909	0.915					
Age: 37 - 49		0.926	0.941	0.948	0.920				
Spouse age: 27-34						0.923	0.949		
Spouse age: 35-47	1.071	1.052						1.228	1.309
Time address: up to 4 m	1.130	1.072	1.075	1.073		1.073			
Time address: 1y - 5y	0.966							0.877	
Time on job: 1.5y -3y		1.039	1.052	1.051					
Time on job: 3y -10y								0.904	

	1	2	*3*	*4*	*5*	*6*	*7*	*8*	*9*
No card insurance		1.466	1.441	1.471	1.533	1.616	1.524	1.297	
No credit insurance		1.201	1.179	1.198	1.276	1.323	1.309		
Credit ins 8		1.260	1.239	1.255	1.368	1.448	1.342		
Product: computers		1.237	1.222	1.241	1.322	1.391	1.360		
Product: video		1.316	1.290	1.304	1.398	1.363	1.311	1.282	
Product: TV		1.465	1.355	1.479	1.674	1.553	1.448	1.861	1.646
Product: furniture		1.300	1.145	1.214	1.379	1.271	1.197	1.508	
Product: cycles		1.245	1.104	1.147	1.228	1.174	1.124	1.500	
Price: under 400 NLG		1.126			1.159			1.468	
Price: 401-800 NLG					1.099	1.094		1.346	
Price: 801-1600 NLG		0.770	0.801	0.790	0.856	0.887	0.813	0.596	0.706
Price: 1601-2200 NLG		1.088			1.070			1.317	
Price: 2201-3000 NLG								1.263	1.314
Agreement*pay date: budget*01		3.688	3.799	3.883	3.424	4.015	3.304	4.925	1.897
Agreement*pay date: budget*other					0.787	0.677	0.726	1.964	0.489
Agreement*pay date: defer		2.750	2.688	2.797	2.927	2.586	2.517	3.653	3.197
ATS $_{t-1}$: over credit limit			0.704	0.452	0.172	0.779	0.801	0.250	0.077
ATS $_{t-1}$: 0-500 NLG			0.835			1.350	1.222	0.645	0.331
ATS $_{t-1}$: 501-1000 NLG				1.124		1.553	1.362	0.704	
ATS $_{t-1}$: 1000-3000 NLG			1.046	1.076	0.882	1.405	1.273	0.740	
ATS $_{t-2}$: over credit limit					2.899				
ATS $_{t-2}$: 0-500 NLG					2.611				
ATS $_{t-2}$: 501-1000 NLG					1.802				
ATS $_{t-2}$: 1000-3000 NLG					2.229				
1 missed payment						0.041			
2+ missed payments						0.022			
Percent repaid						1.310			

Germany

Variable	Model								
	Per-sonal	Personal, Purchase	Personal, Purchase, ATS_1	Personal, Purchase, ATS_{t-1}	Personal, Purchase, ATS_{t-1}, ATS_{t-2}	From 2nd period-more info	Personal, Purchase, ATS_{t-1} – Goods	Personal, Purchase, ATS_{t-1} -Bads after 2nd p	Personal, Purchase, ATS_{t-1} – Bads before 2 p
	1	2	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Phone given	1.100	1.146	1.125	1.114	1.162			1.231	
Res: Owner	0.962								
Res: Rented house	1.156	1.156	1.150	1.146	1.164	1.159	1.142		
Marital: Widowed	1.116								
Kids: 1-2		1.055	1.042	1.032	1.059	1.039	1.027		
Kids: 3+	1.096	1.162	1.152	1.143	1.159	1.141	1.146	1.175	
Occup: Full-time	0.870							1.257	
Occup: Self-emp	0.811	0.942			0.936	0.931			
Type of business: Service	0.930								
Type of business: Education, healthcare	0.832	0.956	0.957	0.957	0.950	0.962	0.960		1.271
Age: under 21	1.066	0.930					1.121		
Age: 21 - 29	1.102		1.031	1.030	1.057	1.040	1.074		
Age: 30-34	1.074				1.040		1.036		
Age: 40-55			0.970	0.970		0.954	1.000		
Age: 55+		0.927	0.871	0.866	0.936	0.871	0.870		
Spouse age: 40-62	0.938								1.622
Time address: up to 6m		0.957	0.965		0.943				
Time address: 6m - 2y	1.032								
Time on job: missing			1.075	1.080			1.095		
Time on job: up to 2y			1.038	1.042		1.050	1.092		

	1	2	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Time on job: 2-4.5y							1.043		
Time on job: 4.5y -7.5y	0.967						1.000		
Time on job: 7.5y -10y	0.944						1.000		
No card insurance		0.805	0.796	0.788	0.857	0.821	0.757	0.873	
Price: under 750 DEM		1.466	1.344	1.351	1.415	1.411	1.412		0.702
Price: 751-1300 DEM		1.320	1.223	1.232	1.278	1.283	1.273		0.752
Price: 1301-1800 DEM		1.177	1.120	1.133	1.158	1.155	1.142		
Price: 1801-2300 DEM		1.079	1.047	1.058	1.075	1.072	1.061		
Product: kitchen		0.876	0.874	0.870	0.862	0.850	0.836		
Product: computers		0.948	0.946	0.939	0.924	0.923	0.906		0.830
Product: household							0.933		
Product: video		1.061	1.053	1.051				1.225	1.357
Pay date: 01		3.999	4.019	4.037	4.720	5.154	4.560	2.422	2.587
Pay date: 08		1.154	1.157	1.152			1.129	1.350	
Agreem: budget		0.805	0.827	0.825	0.719	0.780	0.873	0.618	0.053
ATS $_{t-1}$: over credit limit			0.623	0.131	0.016	0.183	0.322	0.050	0.064
ATS $_{t-1}$: 0-250 DEM			0.706	0.566	0.314	0.719	0.650	0.407	0.204
ATS $_{t-1}$: 251-500 DEM			0.900	0.836	0.606		0.874	0.782	0.603
ATS $_{t-1}$: 501-1500 DEM						1.071		0.913	
ATS $_{t-2}$: over credit limit					1.390				
ATS $_{t-2}$: 0-250 DEM					1.805				
ATS $_{t-2}$: 251-500 DEM					1.187				
1 missed payment						0.016			
2+ missed payments						0.010			
Percent repaid						2.565			

Generic

Variable	Model							
	Personal	Personal, Purchase	Personal, Purchase, ATS_1	Personal, Purchase, ATS_{t-1}	From 2nd period- more info	Personal, Purchase, ATS_{t-1} – Goods	Personal, Purchase, ATS_{t-1} -Bads after 2nd p	Personal, Purchase, ATS_{t-1} – Bads before 2 p
	1	2	*3*	*4*	*6*	*7*	*8*	*9*
Employer's phone given	0.939	0.937	0.921	0.923		0.918	0.869	
Phone given	1.056	1.106	1.088	1.077	1.100		1.308	1.648
Kids: 0	0.914	0.869	0.873	0.868	0.817	0.843	0.852	1.412
Kids: 1-2	0.906	0.857	0.862	0.857	0.806	0.827		1.515
Marital: Married								1.252
Marital: Single, divorced		0.938	0.944	0.954	0.945			
Res: Owner	0.943				0.904			
Res: Rented house	1.038	1.081	1.093	1.093	1.077	1.111		
Occup: Full-time	0.939	0.914	0.914	0.940	0.928	0.932		
Occup: Self-emp	0.769	0.797	0.800	0.819	0.840	0.841	0.763	
Occup: Retired, housewife	0.915	0.865	0.865	0.852	0.865	0.860		
Type of business: Building	0.812	0.960	0.938	0.872	0.952	0.952		0.734
Type of business: Healthcare	0.875			0.909	0.895		0.893	
Type of business: Industry	0.878			0.923	0.929			
Type of business: Shop	0.947			0.949				
Age: under 22		0.947						
Spouse age: no spouse		0.956	0.928	0.934		0.958		
Spouse age: under 26							1.180	
Time address: up to 6m	1.047							
Time address: 17y+	0.935		0.959	0.957	0.947	0.942		
Time on job: 4-14y						0.957		
Time on job: 14+		0.958	0.961	0.957	0.936	0.921		
Time on job: Allowance								1.447

	1	2	*3*	*4*	*6*	*7*	*8*	*9*
No card insurance		0.773	0.801	0.776	0.849	0.814	0.657	
Card insurance		0.846	0.896	0.870	0.899	0.908	0.773	
No credit insurance		1.060		1.048	1.075			0.747
Product: heater, TV		0.937	0.929	0.918	0.917	0.959	0.919	
Product: hifi radio						1.060		
Product: kitchen items						1.049		
Product: card app, video		1.071	1.062	1.075	1.075	1.151		
Price: 0		1.472	1.321	1.387	1.478	1.450	1.239	1.377
Price: 1-180 EUR		1.644	1.441	1.409	1.491	1.461	1.374	1.514
Price: 181-450 EUR		1.490	1.289	1.326	1.423	1.390	1.137	0.848
Price: 451-650 EUR		1.356	1.210	1.247	1.280	1.281	1.125	
Price: 651-800 EUR		1.243	1.130	1.166	1.210	1.198		
Price: 801-1100 EUR		1.136	1.090	1.102	1.128	1.128		
Pay date: 01		2.330	2.331	2.427	2.929	2.520	2.003	1.575
Pay date: 08		1.318	1.312	1.352	1.264	1.351	1.305	
Agreem: budget phone prop		1.803	1.835	1.649	1.649	1.488	2.048	
Agreem: budget other		1.563	1.698	1.535	1.522	1.316	1.972	
Agreem: deferred		1.615	1.566	1.426	1.336	1.264	2.048	3.888
ATS: over credit limit			0.572	0.298	0.402	0.458	0.224	0.097
ATS: 0-150 EUR			0.845	0.834	0.843	0.934	0.738	0.270
ATS: 151-500 EUR				0.969				
ATS: 501-800 EUR			1.049		1.036			
ATS: 801-1200 EUR							1.124	
2+ missed payments					0.001			
Percent repaid					2.098			